[redacted]

THIS PAGE BLANK (USPTO)

PCT WORLD INTELLECTUAL PROPERTY ORGANIZA
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C07K 13/00, C07H 17/00, C12P 21/06, C12N 15/00, 1/20, G01N 33/566, 33/563, A61K 37/00

A1

(11) International Publication Number: WO 94/24158

(43) International Publication Date: 27 October 1994 (27.10.94)

(21) International Application Number: PCT/US94/04132

(22) International Filing Date: 14 April 1994 (14.04.94)

(30) Priority Data:
048,164 14 April 1993 (14.04.93) US

(60) Parent Application or Grant
(63) Related by Continuation
US 048,164 (CIP)
Filed on 14 April 1993 (14.04.93)

(71) Applicants *(for all designated States except US)*: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 300 Lakeside Drive, Oakland, CA 94612 (US). YISSUM RESEARCH DEVELOPMENT COMPANY OF THE HEBREW UNIVERSITY OF JERUSALEM [IL/IL]; 46 Jabotinsky Street, 91042 Jerusalem (IL).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: SHOSEYOV, Oded [IL/IL]; 5 Haerez Street, Karme Yosef D.N., 72910 Shimshon (IL). SHPIEGL, Itai [IL/IL]; Kibbutz Amiad, 12335 North Galilea (IL). GOLDSTEIN, Marc, A. [US/US]; 1959 Lake Boulevard, #227, Davis, CA 95616 (US). DOI, Roy, H. [US/US]; 1520 Lemon Lane, Davis, CA 95616 (US).

(74) Agents: BALDWIN, Geraldine, F. et al.; Pennie & Edmonds, 1155 Avenue of the Americas, New York, NY 10036 (US).

(81) Designated States: AU, BB, BG, BR, BY, CA, CN, CZ, FI, GE, HU, JP, KG, KR, KZ, LK, LV, MD, MG, MN, MW, NO, NZ, PL, RO, RU, SD, SI, SK, TJ, UA, US, UZ, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: CELLULOSE BINDING DOMAIN

(57) Abstract

A cellulose binding domain (CBD) having a high affinity for crystalline cellulose and chitin is disclosed, along with methods for the molecular cloning and recombinant production thereof. Fusion products comprising the CBD and a second protein are likewise described. A wide range of applications are contemplated for both the CBD and the fusion products, including drug delivery, affinity separations, and diagnostic techniques.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

CELLULOSE BINDING DOMAIN

This application is a Continuation-In-Part of United States Application Serial No. 08/048,164, filed April 14, 1993.

1. INTRODUCTION

The present invention relates to a cellulose binding domain (CBD) that binds to water-insoluble forms of cellulose and chitin, including crystalline forms, with a remarkably high affinity and in a reversible manner. The CBD of the present invention, which has been isolated and which is substantially free of other proteins with which the CBD is naturally associated, finds use, for example, in the bio-immobilization of a wide variety of substances, especially biologically active molecules, to cellulose. Fusion products comprising the CBD and a second protein of interest are also disclosed, including applications in methods for their preparation. Such fusion proteins enjoy a wide range of useful applications, including applications in separation, purification and diagnostic methods.

2. BACKGROUND OF THE INVENTION

2.1. Immobilization of Proteins

Immobilization of chemical substances, including biologically active proteins, is of great importance to industry. Various methods of immobilization have been developed in recent years. However, most of these methods require chemical modification of a solid matrix. Typically, these modifications require the covalent attachment of a ligand to the matrix resulting, in many cases, in loss

of activity of the ligand as well as the inclusion of toxic organic compounds that must be removed before the matrix can be used in medicine or food processing. A typical example of a widely used product is Protein A-Sepharose. This highly expensive product is used for the purification of IgG by affinity chromatography, as well as for many diagnostic protocols.

The use of chimeric proteins, *i.e.*, those that contain a functional domain (catalytic or otherwise) together with a binding domain, is relatively new but has already proven to be very useful, especially in protein purification methods. For example, the glutathione S-transferase gene fusion system is designed to express a gene of interest fused to the C-terminal of glutathione S-transferase. The recombinant protein is purified by affinity chromatography using glutathione-Sepharose column.

Another example of a chimeric protein having a functional domain and a binding domain is the Protein-A gene fusion vector which has been designed to permit a high level of expression of fusion proteins in both *E. coli* and *Staphylococcus aureus* cells. B. Nilsson, *et al.* (1985) EMBO J. 4(4):1075-1080. The IgG binding domain of Protein A provides a rapid purification method of the fusion protein using IgG-Sepharose columns. Similar systems have been developed based on beta-galactosidase fusion proteins purified on IPTG-Sepharose or metal chelate chromatography or histidine hexamer fusion proteins using Ni-resin columns. All these methods use expensive matrices such as Sepharose, acrylic beads, or glass beads that require costly chemical modifications and in many cases, the use of highly toxic compounds. Cellulose, on the other hand, has

excellent physical properties and is inexpensive, thus providing an attractive solid matrix for protein immobilization.

2.2. Cellulose

Cellulose has been very useful for immobilization of endo-beta-glucosidase, an enzyme that is used for wine and fruit juice treatments (Shoseyov *et al.* J. Agric. Food Chem. 38:1387-1390 [1990].). However, immobilization requires the chemical modification of the cellulose and results in a fluffy compressible material that is not suitable for applications involving packed columns.

A cellulose substrate, to which a ligand could be bound without resort to chemical modification (*e.g.*, "bioimmobilized"), would possess the advantage that the resulting solid matrix is natural and non-toxic (having required no "dirty chemical modifications"). It would be further advantageous if the resulting solid matrix retained its physical properties, as well as its relatively low price. At present, cellulose prices are 100-500 fold lower than those of glutathione-Sepharose and IPTG-Sepharose, making cellulose an attractive, inexpensive matrix that can be used safely in food and pharmaceutical industries. Recently, Greenwood et. al. (FEBS Lett. 224(1): 127-131 [1989].) fused the cellulose binding region of *Cellulomonas fimi* endoglucanase to the enzyme alkaline phosphatase. The recombinant fusion protein retained both its phosphatase activity and the ability to bind to cellulose. *See*, also U.S. Patent No. 5,137,819 granted to Kilburn *et al.*, incorporated by reference herein.

Unfortunately, the *Cellulomonas fimi* cellulose binding region exhibits a relatively low

affinity to cellulose. For instance, more than 30% of the fusion protein is washed off by 50 mM Tris-HCl (pH 7.5) in 0.5 M NaCl. A second disadvantage of the *C. fimi* cellulose binding region is that the cellulose fibers are disrupted upon binding to the *C. fimi* binding region (Din *et al.*, Bio/Technology 9:1096-1098 [1991]). Therefore, even though the *C. fimi* cellulose binding region may not exhibit direct cellulase activity, this disruption of the fibers of the cellulose substrate, which is tantamount to a physical change in the morphology of the solid matrix, is equally problematic and undesirable. (*see*, further, the discussion in Section 7.3, below.)

2.3. Cellulase

Many cellulases and other hydrolytic enzymes, such as chitinases, have high affinities for their substrates (Shoseyov, *et al.*, PNAS USA 87:2192-2195 [1990]; Cabib, Methods Enzymol. 161, pp. 460-462 [1988]). It has previously been shown that strong binding between crystalline cellulose and the cellulase is directly related to that cellulase's ability to degrade crystalline cellulose (Klyosov, Biochemistry 29:10577-10585 [1990]), whereas strong binding is not necessary for that cellulase's ability to degrade amorphous cellulose.

Shoseyov, *et al.*, (PNAS USA 87:2192-2195 [1990]) report purification of the cellulase "complex" from *Clostridium cellulovorans*. This cellulase "complex" exhibits cellulase activity against crystalline cellulose, as well as carboxymethylcellulose, and is a large protein complex consisting of several different polypeptides. It has been found that a large (ca. 200 kD) major cellulose binding protein (CbpA), with no apparent enzyme

activity, must participate with the catalytic enzyme in order for the catalytic enzyme to breakdown crystalline cellulose. However, no such participation by the CbpA is necessary for the enzymatic degradation of amorphous cellulose. Shoseyov, *et al.*, (PNAS USA 89:3483-3487 [1992]) report the cloning and DNA sequencing of the gene for CbpA, including a description of four separate "putative" cellulose binding domains within the CbpA.

2.4. Heat Shock Protein (HSP)

Heat shock proteins (HSP) are induced in prokaryotic and Eukaryotic species under various conditions of stress. The HSP's are grouped into families of homologous proteins based on their molecular masses.

The 60 kD HSP (hsp 60) family, which retained a uniquely high level of sequence conservation during evaluation is a focus of interest as a potential antigen in autoimmune diseases. W. Van Eden (1991) Immunol. Rev. 121:1; W. Van Eden, Thole R. Van der Zee, A. Noordzij, J.D.A. Van Einbden, E.J. Hensen, and I.R. Cohen (1988) Nature 331:171; M.B.J. Billingham S. Carney, R. Bulter, and M.J. Colston (1990) J. Exp. Med. 171:339.

There is experimental evidence that response to hsp 60 is subject to regulatory T cell control. It has been suggested that T cell reactivity is directed, at least partly, against the mammalian endogenous HSP. B. Herman *et al.* (1991) Eur. J. Immunol. 21:2139.

The genes coding for hsp 60 protein family have been cloned and sequenced from a number of different species including *Mycobacterium tuberculosis, Microbacterium leprae, Mycobacterium bovis BCG, E. coli*, chinese hamster rat mouse and

human cells. V. Mera *et al.* (1986) Proc. Nat. Acad. Sci. USA 83:7013; S. Jindal *et al.* (1989) Mol. Cell. Biol. 9:2279; O.J. Picketts *et al.* (1989) J. Biol. Chem. 264:12001; T.M. Shinnick (1987) J. Bacteriol. 169:1080; T.M. Shinnick *et al.* (1988) Infect. Immun. 56:446; T.J. Venner and R.S. Gupta (1990) Nucl. Acids. Res. 18:5309.

As described *infra*, according to the present invention, the heat shock protein (HSP), portions thereof, anti-HSP antibodies or portions thereof, can be used in CBD-fusion products.

Immune response to Mycobacterial and human hsp 60 has been implicated in the development of autoimmune diabetes in human and animal models.

Besides being an immunodominant antigen, the hsp 60 family of proteins in various systems has been shown to perform a "molecular chaperone" role in the proper folding of newly synthesized polypeptide chains and, in some cases, their assembly into oligomeric protein complexes. S.M. Hemmingsen *et al.* Nature 333:330; R.J. Ellis (1990) Sem. Cell Biol. 1:1-9.

In a model system in mice for insulin dependent diabetes mellitus (IDDM) - T lymphocytes responding to the hsp 60 antigen are detectable at the onset of insulitis and it is likely that these T lymphocytes can also recognize the Beta cell hsp 65 cross reactive antigen. D. Elias *et al.* (1990) Proc. Nat. Acad. Sci. USA 87:1576-1580.

Thus, there remains an unfulfilled need to discover a cellulose binding protein that exhibits a high but reversible affinity for cellulose, particularly crystalline cellulose, but which manifests neither cellulase activity nor an ability to disrupt the fibers of the crystalline cellulose

substrate (*i.e.*, does not exhibit an "amorphogenic" effect).

3. SUMMARY OF THE INVENTION

The inventors describe herein the identification, molecular cloning and cellulose binding characteristics of a novel cellulose binding domain (CBD) protein. Also, disclosed is the construction of an expression system for the production of a fusion product comprising the CBD and a second protein of interest.

The invention relates to the discovery that the CBD is able to function independently of the other proteins or polypeptides with which it is naturally associated. Moreover, it has been discovered unexpectedly that the CBD has a high affinity for crystalline cellulose and chitin, having a K_d ranging from about 1.5μM to about 0.5μM, preferably ranging from about 1.4μM to about 0.8μM. In particular, with various samples of crystalline cellulose, the CBD exhibits a K_d of about 1.2μM or less. The CBD can be further characterized in that it possesses virtually no cellulase activity and, quite surprisingly, the CBD exhibits no morphology-altering characteristics (*i.e.*, no amorphogenic effects). It has also been discovered that CBD-fusion products comprising CBD and a second protein retain the avid binding capacity of the CBD to cellulose.

The invention is also related to the discovery that the CBD demonstrates absolute binding to cellulose over a wide range of pH and under different buffering conditions, that large quantities of CBD bind to crystalline cellulose, and that exposure to water fails to release CBD from cellulose. In stark contrast, the major CbpA protein, as well as

the binding region from *C. fimi*, are readily dissociated from cellulose on exposure to water. Indeed, exposure to denaturing solutions, such as 6 M guanidine-HCl, 6 M urea or nonionic surfactants, is required to release the CBD from cellulose. Thus, the CBD protein functions and behaves quite independently of the rest of the other proteins with which it is naturally associated in its binding to cellulose.

Thus, the present invention provides an isolated CBD protein capable of binding cellulose with high affinity and which is substantially free of other proteins with which it is naturally associated.

In one embodiment of the present invention, a CBD protein comprises the amino acid sequence shown in Figure 1. In another embodiment of the present invention, the CBD protein comprises an amino acid sequence having at least 70% homology, preferably 80% homology, to the amino acid sequence disclosed in Figure 1. In another aspect of the present invention, a nucleic acid is contemplated having at least 60% homology, preferably 70-80% homology, to the nucleic acid sequence depicted in Figure 1.

In yet another embodiment of the present invention, the CBD protein has a high binding affinity to cellulose. More preferably, the CBD of the present invention has a K_d of about 1.5μM or less, most preferably 1.0μM or less.

In yet another embodiment, the present invention provides a CBD fusion protein comprised of a CBD protein capable of binding cellulose with high affinity and a second protein wherein, when said CBD protein is one which occurs in nature, said CBD protein is substantially free of other proteins with which it is naturally associated. In a particular embodiment of the present invention, the second

protein is Protein A. In another embodiment of the present invention, the second protein is an HSP protein. In yet another embodiment of the present invention, the second protein fused to the CBD may be comprised of two or more polypeptide regions. For example, the CBD may be fused to the variable light chain (V_L) and the variable heavy chain (V_H) of an antibody or functional portions thereof.

A further embodiment of the present invention provides a method for the production of a CBD or a CBD fusion product. An exemplary procedure, which can be applied either to CBD alone or to a CBD fusion product, but which is recited herein for a CBD fusion product, may comprise the following steps: providing nucleic acid encoding the CBD fusion product wherein said CBD fusion product is comprised of a CBD and a second protein, said CBD being capable of binding cellulose or chitin with high affinity and being substantially free of other proteins with which it is naturally associated; transfecting a host cell with the nucleic acid or using an equivalent means for introducing the nucleic acid into the host cell; and (2) culturing the transformed host cell under conditions suitable for expression of the CBD fusion protein. In the above method, the second protein may be further comprised of an N-terminal amino acid and a C-terminal amino acid. Transfection of the host cell can be effected in a number of ways well known to those of ordinary skill in the art, including, but not limited to, electroporation, injection, calcium chloride precipitation and retroviral introduction. Furthermore, the nucleic acid can be either integrated with the genome of the host cell or not.

In a further aspect of the present invention, a method of the purification of a CBD

fusion product is provided comprising contacting a mixture comprising a recombinant CBD fusion product with an effective amount of cellulose under conditions suitable for the formation of an insoluble binding complex comprising cellulose and the recombinant fusion product; isolating the insoluble cellulose-CBD fusion product binding complex from the mixtures; and recovering the CBD fusion protein from the cellulose-CBD fusion product binding complex.

In a particular embodiment of the present invention, the method of purifying a CBD fusion protein of the present invention further comprises providing nucleic acid encoding a cleavage site upstream of the N-terminal amino acid of the second protein of the CBD fusion product.

In another embodiment of the present invention, the method for purifying a CBD fusion protein further comprises providing nucleic acid encoding a cleavage site downstream of the C-terminal amino acid of the second protein of the CBD fusion protein.

In yet another embodiment of the present invention the method for purifying a CBD fusion protein of the present invention further comprises providing nucleic acid encoding a first cleavage site upstream of the N-terminal amino acid of the second protein and a second cleavage site downstream of the C-terminal amino acid of the second protein of the CBD fusion protein.

Similarly, CBD is purified by contacting a mixture comprising CBD with an effective amount of cellulose. Isolation of the resulting insoluble CBD-cellulose binding complex followed by treatment of the binding complex with a releasing reagent provides the purified CBD.

Another aspect of the present invention provides an isolated nucleic acid encoding a CBD protein which is capable of binding cellulose with high affinity and which is substantially free of other nucleic acid with which it is naturally associated. An isolated nucleic acid encoding a protein of the present invention may be useful as a probe in screening cDNA or genomic libraries for sequences having homology to the *cbd* gene.

In further embodiments, the present invention provides for a host cell comprised of nucleic acid encoding a CBD of the present invention.

An additional aspect of the present invention relates to a diagnostic kit for the detection of a substance of interest comprising: (a) a CBD fusion product comprising (i) a CBD capable of binding to cellulose with high affinity and substantially free of other proteins with which it is naturally associated, and (ii) a second protein capable of binding a substance of interest; (b) a detectable label; and (c) cellulose. In such a diagnostic kit, the cellulose can be replaced by chitin. In specific embodiments of the present invention, the second protein of the CBD fusion product can be Protein A, HSP protein, HSP antibody, HSP-related protein, peptide or antigenic portion thereof. The term "peptide" is meant to include molecules comprising 2-20 amino acids. The CBD fusion product of the disclosed diagnostic kit may further comprise a ligand affinity bound to the second protein, such ligand capable of binding a substance of interest. By "ligand" is meant any molecule that is able to bind a second molecule by any non-covalent means. For example, a primary IgG can be affinity bound to Protein A fused to CBD. The IgG may then

serve as a ligand for a particular protein, peptide or hormone.

In another aspect of the present invention, an immunoassay method of detecting the presence of a substance of interest in a test sample is disclosed comprising: (a) incubating a test sample, which may contain a substance of interest, with a sufficient amount of a CBD fusion product comprising (i) a CBD capable of binding to cellulose with high affinity and substantially free of other proteins with which it is naturally associated, and (ii) a second protein capable of binding the substance of interest, under conditions that allow for the binding of the substance of interest to the second protein of the CBD fusion product; (b) adding an amount of cellulose effective to bind the amount of the CBD fusion product used in step (a) to provide an insoluble cellulose-CBD fusion product binding complex; (c) separating the insoluble cellulose-CBD fusion product binding complex from unbound components; (d) incubating the insoluble cellulose-CBD fusion product binding complex with a sufficient amount of a detectable label, the label capable of binding to the substance of interest; and (e) separating the insoluble cellulose-CBD fusion product binding complex of step (d) from unbound components and determining the presence or absence of the label, to provide an indication of the presence or absence of the substance of interest in the test sample.

Another method is also disclosed for the detection of a substance of interest in a test sample comprising: (a) contacting a test sample, which may contain a substance of interest, with an insoluble matrix capable of immobilizing the substance of interest; (b) incubating the insoluble matrix with a

sufficient amount of a CBD fusion product comprising (i) a CBD capable of binding to cellulose with high affinity and substantially free of other proteins with which it is naturally associated, and (ii) a second protein capable of binding the immobilized substance of interest, under conditions that allow for the binding of the immobilized substance of interest to the second protein of the CBD fusion product; (c) separating the insoluble matrix of step (b) from unbound components; (d) incubating the insoluble matrix of step (c) with a detectable label capable of binding the substance of interest or the CBD fusion product under conditions that allow for the binding of the label to the substance of interest or the CBD fusion product; and (e) separating the insoluble matrix of step (d) from unbound components and determining the presence or absence of the label, to provide an indication of the presence or absence of the substance of interest in the test sample.

This method may further comprise (i) contacting the insoluble matrix of step (c) with a sufficient amount of cellulose under conditions that allow for the binding of the cellulose to the CBD fusion product to form a cellulose-CBD fusion product binding complex, and (ii) separating the insoluble matrix of step (i) from unbound components, including unbound cellulose. This method can use a label that is capable of binding the substance of interest or the cellulose-CBD fusion protein binding complex, in particular, the cellulose of the cellulose-CBD fusion protein binding complex. The test sample may be a bodily fluid, including, but not limited to, blood, urine, semen, saliva, mucus, tears, vaginal secretions, and the like. As usual, the cellulose can

be replaced by chitin. Also, the insoluble matrix may be an electrophoresis gel blot.

In a specific embodiment of the present invention, the method is designed for the detection of a protein or peptide; thus, the second protein of the CBD fusion product may be an antibody against the protein or peptide. The antibody may be a monoclonal antibody or a polyclonal antibody. Alternatively, the CBD of the present invention may be conjugated to the antibody of interest either directly or through a linker moiety.

The substance of interest may also comprise a biotinylated probe bound to a protein, peptide, hormone, nucleic acid or other probe-targetable molecule. In this case, the preferred second protein is streptavidin. Where the label includes an enzyme, the method further comprises adding a sufficient amount of a substrate for the enzyme, which substrate is converted by the enzyme to a detectable compound.

The assay may also be carried out in a competitive mode. Hence, the above-described method may be modified such that step (d) is performed by incubating the insoluble cellulose-CBD fusion product binding complex with a sufficient amount of a detectable label comprising a labeled substance of interest, the label capable of binding to any second protein of the CBD fusion product which remains unbound to the substance of interest, and in which step (e) is performed by separating the insoluble cellulose-CBD fusion product binding complex of step (d) from unbound components and comparing the signal observed from the test sample relative to the signal observed from a control sample.

In a preferred embodiment of the present invention, the CBD fusion product is included in a dip

stick. Hence, it is also an object of the present invention to provide a dip stick useful in detecting a substance of interest in a test sample comprising a CBD fusion product, the CBD fusion product comprising (i) a CBD capable of binding to cellulose with high affinity and substantially free of other proteins with which it is naturally associated, and (ii) a second protein capable of binding a substance of interest. Preferably, the second protein is selected from the group consisting of Protein A, HSP protein, HSP antibody, cross-reactive HSP-related protein or peptide or an antigenic portion thereof, an enzyme, hormone, antigen, and antibody.

Likewise, it is also an object of the present invention to provide a signal amplification system comprising: (a) a first CBD fusion product comprising (i) a CBD capable of binding to cellulose with high affinity and substantially free of other proteins with which it is naturally associated, and (ii) a second protein capable of binding a chimeric probe, the probe further capable of binding a substance of interest; and (b) a second CBD fusion product comprising (i) a CBD capable of binding to cellulose with high affinity and substantially free of other proteins with which it is naturally associated, and (ii) an enzyme capable of acting on a substrate to produce a detectable compound.

In another embodiment, a signal amplification system is provided which comprises: (a) a CBD fusion product comprising (i) a CBD capable of binding to cellulose with high affinity and substantially free of other proteins with which it is naturally associated, and (ii) a second protein capable of binding a chimeric probe, the probe further capable of binding a substance of interest; and (b)

labeled CBD, the CBD retaining its capacity to bind to cellulose with high affinity and substantially free of other proteins with which it is naturally associated.

These signal amplification systems may include the chimeric probe and may further include a cellulose matrix, preferably a pebble-milled cellulose.

These signal amplification systems may include systems designed to identify nucleic acid from cells or tissue and systems designed to extend the length of identified nucleic acid in either direction which employ oligonucleotide-directed thermocyclic DNA amplification of the missing sequences, wherein a combination of a DNA-specific primer and a degenerate, vector-specific, or oligo-dT-binding second oligonucleotide can be used to prime specific synthesis.

Finally, it is a further object of the present invention to provide a drug delivery system comprising CBD associated with a drug, the CBD retaining its capacity to bind to cellulose with high affinity and substantially free of other proteins with which it is naturally associated. In such a drug delivery system, the drug is conjugated to the CBD either directly or through a linker moiety. Many methods of conjugation exist and are known in the art. For example, acyl activation agents exist, such as cyclohexylcarbodiimide, which can be used to form amide or ester bonds. Thus, a drug having a nucleophilic group, such as amino or hydroxy may be attached to the carboxy terminal end of CBD. In one embodiment such a drug delivery system can be a slow or sustained drug delivery system wherein the drug of interest is slowly released from the CBD bound to cellulose.

In one embodiment, the drug to be delivered is an antifungal agent. Preferred agents, include, but are not limited to, Amphotericin B, Nystatin, and Undecylenic Acid. The drug may generally be an imidazole, such as Clotrimazole. It is contemplated that such a drug delivery system can be incorporated into a composition that can be administered parenterally, orally, topically or by inhalation. In particular, routes of administration include, but are not limited to, intranasal, opthalmic or intravaginal. Furthermore, the composition may be in the form of a solid, gel, liquid or aerosol.

The drug delivery system described herein are useful for the delivery of a drug to an infectious or disease-causing agent, such as a yeast or fungal agent whose cellular membrane contains a cellulosic or chitinic substance. Examples of such agents, include, but are not limited to, *Aspergillus fumigatus*, a member of the genus *Candida* or *Monilia*, or an epidermatocyte. By employing the drug delivery system of the present invention, it is anticipated that the dosages required for effective treatment of the disease or infection will be much reduced, thus, improving the effectiveness of the antifungal or antimycosal drugs, which are typically also quite toxic. It is hoped that side effects are, thus, also minimized or, even better, eliminated.

Yet another aspect of the present invention relates to the ability of CBD to modify the growth of plant tissues. Such ability finds application in the agricultural industry, for example, where it may be desirable to promote pollen tube growth and/or root growth. In addition, since CBD in high concentration has the ability to inhibit root growth it is useful to prevent the devlopment of undesirable plants, for

example, "weed" plants. In another example, the CBD finds application in enhancing growth of fungi, such as mushrooms, wherein the CBD binds chitin and stimulates the growth and/or development of developing basidiocarps or fruiting bodies.

Yet another aspect of the present invention relates to the ability of CBD to bind to chitin of insect exoskeletons and other insect parts, including the mid-gut. Such ability finds application in the use of CBD as a bio-pesticide wherein the CBD is linked to a microorganism useful for controlling insect pests, such as *Bacillus thuringiensis* (BT) or other microorganism expressing the BT toxin gene, or such as a chitinase secreting microorganism or fungus. The CBD functions to bind the microorganism to the chitinous portions of the insect where the microbially produced toxin or enzyme leads to death of the insect pest. Microorganisms which display chitinase activity include bacterial strains from the genera Enterobacter and Streptomyces, and fungal strains from the genera Aspergillus, Penicillium and Trichoderma, etc. In one mode of this embodiment of the invention, the CBD-microorganism complex can be applied directly to plant parts suspected or at risk of infestation by insect pests or, alternatively, the CBD-microorganism complex can be immobilized on bait stations which are set out in the infested fields or on plants infested with insect pests. In another alternative mode of this embodiment of the invention, a CBD-microorganism complex is applied to wood or cellulose-based products to prevent or control infestation by wood destroying insect pests. For example, a CBD-microorganism complex is applied to telephone poles or wood or cellulose-based construction components where the CBD component binds tightly to invading insects and the

microorganism of the complex is effective to control the insect pests.

Still another aspect of the present invention relates to the ability of CBD to bind to chitin found in fungi, including fungi which are useful for bioremediation or degradation of toxic environmental pollutants. According to one mode of this embodiment, a CBD fusion protein is used in which the second component is an enzyme which degrades the environmental pollutant. The CBD-enzyme fusion protein can be immobilized on a cellulose substrate and the device used to remove the pollutant from the environment by enzymatic degradation, for example, by placing the device in a stream of surface or underground water. According to an alternative mode of this embodiment of the invention, the CBD is bound to a fungus which secretes an enzyme which is useful for biomediation, *i.e.*, which degrades an environmental pollutant such as a pesticide. The CBD-fungus complex is then immobilized on a cellulose substrate and provides a useful device to enzymatically degrade an environmental pollutant such as a pesticide including DDT or such as wood containing products left over from logging or saw milling operations. In one illustrative example, CBD is complexed with a white rot fungus, for example *Phanerochaete chrysosporium*, which produces lignin peroxidase, manganese peroxidase and degrades the pesticide DDT. The CBD-fungus complex is useful for bioremediation, for example, to remove a pesticide such as DDT from the environment.

4. DESCRIPTION OF THE FIGURES

Figure 1. Nucleotide (top, orig.; second line, complement) and deduced amino acid sequence of CBD.

Figure 2. Preparation and cloning of the gene fragment encoding for CBD.

A. Analysis of the primary structure of CbpA, which contains an N-terminus signal peptide, unique CBD region, 4 hydrophilic repeats (white arrows), and 9 hydrophobic repeats (black arrows).

B. PCR primer placement along the *cbp*A gene. Included for clarity are the primer sequences and the *cbp*A DNA sequence of the CBD flanking regions. The PCR product contains *Nco*I and *Bam*HI sites, underlined. Also note that the ATG start codon for the gene fragment is located within the *Nco*I site, and the TAG stop codon is adjacent to the *Bam*HI site.

C. Schematic of pET-CBD, containing the CBD gene fragment cloned into the pET-8c vector. The vector contains the necessary transcriptional and translational signals for inducible CBD production.

Figure 3. Expression and purification of the CBD protein.

Whole cell proteins from cells harboring pET-8c (*lane 2*), whole cell proteins from cells harboring pET-CBD (*lane 3*), cytosolic fraction from lysed pET-CBD cells (*lane 4*), Guanidine HCl-solubilized membrane/ inclusion body fraction from lysed pET-CBD cells (*lane 5*), final PC buffer wash of Avicel® pellet (*lane 6*), and purified CBD protein (*lane 7*) were loaded on a 15% acrylamide gel. Each lane was loaded with 0.005% of the total protein of each fraction, except lane 6 which is a 10-fold concentrate. Prestained molecular mass markers (*lanes 1,8*) have mobilities of approximately 2.6, 5, 12.7, 18.1, 29, and 44 kDa.

Figure 4. Time course of CBD-Avicel® binding. CBD (2.0 μmol total protein) and Avicel® (1 mg/ml) were equilibrated as described in Section 7, below, except that a larger total volume was used to provide samples taken at various time points. Each time point sample was washed and assayed as described in Section 7. *see*, also, Table I of Figure 9, below.

Figure 5. Double reciprocal plot of CBD binding to Avicel.® 0.5 mg Avicel® is represented by closed squares (B), 1 mg Avicel® by closed circles (J), and 2 mg by closed triangles (H). Inset: PC_{max} versus the amount of Avicel® used. The assay volume was 1.0 ml.

Figure 6. Scatchard Plot CBD binding to Avicel®. The [PC]/[P] vs. PC for 3 amounts of Avicel® are shown. The [PC]/[P] is expressed as a dimensionless ratio, and the [PC] is shown in μM. 1 mg Avicel® is represented by closed circles, 2 mg Avicel® by open circles, and 3 mg by closed squares.

Figure 7. Double reciprocal plot CBD binding to Cellulon®. The incubation mixture contained 0.5 mg Cellulon® per ml.

Figure 8. CBD-ProtA fusion protein binds both Cellulose and IgG. 1M Acetic acid selectively releases CBD-ProtA:IgG "bond", but not the CBD-ProtA:cellulose "bond".

Figure 9. Adsorption of CBD protein to insoluble substrates.

Figure 10. Overexpression and purification of CBD.

Figure 11. Schematic of a signal amplification system in which labeled CBD detects cellulose or chitin to which is bound a CBD fusion product. The substance of interest is bound, in turn, to a chimeric probe.

Figure 12. Schematic of a signal amplification system in which the substance of interest is bound to a ternary CBD fusion product: CBD-ProteinA(IgG).

Figure 13. IgG purification by CBD-ProtA-Cellulose as demonstrated by detection of samples on gel electrophoresis. Lane a is human serum. Lane b are human serum proteins that did not bind to CBD-ProtA-Cellulose. Lane c is the first PBS wash of samples, after binding of the IgG to the ProtA has taken place, showing unbound protein. Lane d is the IgG fraction recovered by acetic acid wash.

Figures 14A and 14B. Demonstration that CBD-treated cellulose fibers remain intact, *i.e.*, CBD has no cellulose disruption amorphogenesis activities. Figure 14 A shows a cotton fiber treated with CBD. Figure 14 B shows untreated cotton fibers (control).

Figure 15. Demonstration of the effect of pH and NaCl concentrations on the binding capacity of CBD-ProtA to cellulose.

Figure 16. Demonstration of CBD binding to *Aspergillus niger*. Lane a is *Aspergillus niger* with CBD. Lane b is *Aspergillus niger* without CBD (control). Lane c is CBD alone.

Figure 17. Demonstration of CBD binding to *Spodoptera littoralis*. Lane a is the mid-gut membrane of *Spodoptera littoralis* without CBD (control). Lane b is the mid-gut membrane of *Spodoptera littoralis* with CBD.

Figure 18. Demonstration of CBD binding to *Heliothis armigera*. Lane a is mid-gut membrane of *Heliothis armigera* without CBD (control). Lane b is mid-gut membrane of *Heliothis armigera* with CBD.

Figure 19. Demonstration of the effect of CBD and BSA (as a control) on pollen tube growth.

Figures 20A and 20B. Demonstration of the effect of CBD on crystalline cell wall components of pollen tubes that were stained with white fluorescent brightener (calcofluor). Figure 20A shows that CBD treated pollen tubes produced non-crystalline pollen tube tip as indicated by the absence of the bright color at the tip zone. Figure 20B shows that a control pollen tube tip contains more crystalline cell wall components as indicated by the bright color.

Figures 21A and 21B. Demonstration of gold-immunolabeling of CBD treated pollen tubes. Figure 21A demonstrates that in pollen treated with CBD, gold-labeling is along the pollen tube but preferably at the tip zone, as indicated by the intensive dark dots at the tip zone. Figure 21B shows the control, *i.e.*, pollen without the CBD treatment.

Figure 22. Demonstration of the effect of CBD on root growth of *Arabidopsis thaliana* seedlings. Different letters indicate statistically significant differences between the root length values ($p \leq 0.05$).

Figure 23. Demonstration of gold-immunolabeling of CBD treated *Arabidopsis thaliana* seedlings. This figure shows that the gold-labeling is restricted to the root and no labeling is shown on the shoot.

Figure 24. Demonstration of *Arabidopsis thaliana* seedling staining with white fluorescent brightener (calcofluor). This figure demonstrates that only the root is stained indicating accessible cellulose.

5. DETAILED DESCRIPTION OF THE INVENTION AND DEFINITIONS

The present invention is directed to the identification of cellulose binding domain (CBD) protein that is capable of binding cellulose with high

affinity and in a reversible manner. The CBD of the present invention may be used, for example, in the bio-immobilization of biologically active molecules to cellulose. The CBD of the present invention may be fused to a second protein to form a CBD fusion protein. The presence of a CBD protein in a CBD fusion protein allows for easy and selective purification of the CBD fusion protein by incubation with cellulose. Examples of second proteins include: Protein A, protein G, streptavidin, avidin, Taq polymerase and other polymerases, alkaline phosphatase, RNase, DNase, various restriction enzymes, peroxidases, glucanases such as endo-1,4-beta glucanase, endo-1,3-beta-glucanase, chitinases, and others, beta and alfa glucosidases, beta and alpha glucoronidases, amylase, transferases such as glucosyl-transferases, phospho-transferases, chloramphenicol-acetyl-transferase, beta-lactamase and other antibiotic modifying and degrading enzymes, luciferase, esterases, lipases, proteases, bacteriocines, antibiotics, enzyme inhibitors, different growth factors, hormones, receptors, membranal proteins, nuclear proteins, transcriptional and translational factors and nucleic acid modifying enzymes. Specifically, the CBD protein may be fused to an antibody or an antigenic determinant to form a CBD fusion product that is useful in diagnostic kits and in immunoassays.

Thus, for example, bodily fluids can be tested for the presence of particular antibodies (*e.g.*, heat shock protein (HSP) antibody) by making use of a CBD and an HSP epitope. Conversely, an HSP protein, a cross-reactive HSP-related protein, or antigenic portions thereof can be detected using a CBD-HSP antibody fusion protein.

The term "CBD" or "CBD protein" or "cellulose binding domain protein" refers to a protein comprising the amino acid sequence shown in Figure 1 and includes functional homologs and functional derivatives thereof, provided that the functional homolog or functional derivative possesses the capability of binding to cellulose with high affinity and in a reversible manner. The CBD of the present invention is provided substantially free of other proteins with which it is naturally associated, for instance, the balance of the major CbpA protein, discussed above. In addition, one or more predetermined amino acid residues in the polypeptide may be substituted, inserted, or deleted, for example, to produce a CBD having improved biological properties, or to vary binding and expression levels. Some of the desired CBD proteins falling within the scope of the present invention may optionally possess covalent or non-covalent modifications of the naturally occurring molecule, including, but not limited to, glycosylation modifications. Through the use of recombinant DNA technology, the CBD proteins of the present invention having residue deletions, substitutions and/or insertions may be prepared by altering the underlying nucleic acid. The modifications or mutations that may be made in the DNA encoding the CBD of the present invention must not alter the reading frame and preferably will not create complementary regions that could produce secondary mRNA structure (*see*, European Patent Publication No. EP 75,444)

The CBD protein of the present invention is one having at least 70% homology to the amino acid sequence shown in Figure 1, preferably, at least 80% homology, more preferably, at least 90% homology, and

most preferably, at least 95% homology. The term "X% homology" is not intended to be limited to sequences having a X% homology over the entire length of the protein. The 70% homology is also intended to include X% homology occurring in identified functional areas within the CBD protein of Figure 1. An example of a functional area would be a defined set of amino acids having the ability to bind cellulose with high affinity and in a reversible manner. Such protein homologs may also be referred to herein as "CBD functional homologs." In one embodiment of the present invention, such a functional area may have about 100 amino acids. In another embodiment of the present invention, such a functional area may have about 50 amino acids. The most desirable CBD protein of the present invention is one comprised of the amino acid sequence shown in Figure 1.

The term "CBD functional derivative" as used herein refers to any "fragment", "variant", "analogue" or "chemical derivative" of the CBD protein amino acid sequence shown in Figure 1 which retains the capability of binding to cellulose with high affinity and in a reversible manner and is preferably between about 2 and about 160 amino acids in length, more preferably between about 25 and about 125 amino acids in length and most preferably between about 50 and about 100 amino acids in length.

The term "fragment" is used to indicate a CBD protein which is derived from the CBD protein shown in Figure 1, and has a naturally occurring sequence. Such a fragment may be produced by proteolytic cleavage of the full-length protein. Alternatively, the fragment is obtained recombinantly by appropriately modifying the DNA sequence encoding the CBD protein to delete one or more amino acids at

one or more sites of the C-terminal, N-terminal, and within the naturally occurring sequence. Fragments of the CBD protein can be screened for the ability to bind cellulose with high affinity and in a reversible manner to determine the identity or utility of a functional derivative.

The term "variant" as used herein is defined as a molecule in which the amino acid sequence, glycosylation pattern, or other feature of a naturally occurring molecule has been modified covalently or noncovalently and is intended to include mutants. Some of the variants falling within this invention possess amino acid substitutions deletions, and/or insertions provided that the final construct possesses the desired ability of binding cellulose with high affinity and in a reversible manner. Amino acid substitutions in the CBD protein may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues involved. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; amino acids with uncharged polar head groups or nonpolar head groups having similar hydrophilicity values include the following: leucine, isoleucine, valine; glycine, alanine; asparagine, glutamine; serine, threonine; phenylalanine, tyrosine. Also included within the definition of variant are those proteins having additional amino acids at one or more sites of the C-terminal, N-terminal, and within the naturally occurring CBD sequence as long as the variant retains the capability of binding cellulose with high affinity and in a reversible manner.

The term "chemical derivative" as used herein refers to a CBD protein produced by chemical modification of naturally occurring or variant CBD protein. Illustrative of an example of a chemical modification would be replacement of H by an alkyl, acyl, or amino group.

The phrase "binding cellulose with high affinity" as used herein refers to the ability of the CBD protein to bind to cellulose with a K_d in μM ranging from about 1.5 to about 0.5, preferably from about 1.4 to about 0.8. More preferably, the high affinity binding refers to the ability of the instant CBD to bind to crystalline cellulose with a K_d of about 1.1 or less, most preferably about 1.0 or less.

The phrase "binding cellulose in a reversible manner" as used herein refers to the ability of the CBD protein to be released from the cellulose-CBD protein binding complex by releasing agents or solutions, such as 6M urea, 6M guanidine-HCl and other denaturing reagents, including nonionic surfactants. Preferably, however, those denaturing reagents are used which allow the released CBD or fusion product to be reconstituted. For example, the CBD may be reconstituted from the treatment with 6M urea or 6M quanidine-Hcl by subjecting the denatured protein to renaturing conditions described in Sections 6.1.4, 7.1.1, 7.2.1, and 8.1.4, below.

The term "CBD fusion protein" as used herein refers to the joining together of at least two proteins, a CBD protein and a second protein. In some embodiments of the present invention, the second protein may be fused or joined to a third protein. In the present invention, examples of second proteins include enzymes, such as nucleic acid modification enzymes, proteases, hormones or hormone precursors,

polypeptides, peptides, antibodies, antigens, antigenic epitopes and variants thereof. In some preferred embodiments of the present invention, the second protein is Protein A; in other preferred embodiments of the present invention, the second protein is an HSP protein. One preferred embodiment of the present invention is a fusion protein comprised of CBD protein, Protein A or anti-HSP recombinant IgG. The CBD fusion protein of the present invention may comprise an enzymatic or chemical cleavage site upstream and preferably adjacent to the N-terminus of the second protein and/or an enzymatic or chemical cleavage site downstream and preferably adjacent the C-terminus of the second protein thereby providing a means for recovering the second protein from the CBD fusion protein through use of a cleaving agent.

The term "CBD fusion protein-cellulose binding complex" as used herein refers to the complex formed when cellulose binds the CBD protein of a CBD fusion protein.

"Nucleic acid" refers to a nucleotide sequence comprising a series of nucleic acids in a 5' to 3' phosphodiester linkage that may be either an RNA or a DNA sequence. If the nucleic acid is DNA, the nucleotide sequence is either single or double stranded. CBD protein encoding nucleic acid is RNA or DNA that encodes a CBD protein capable of binding cellulose with high affinity, is complementary to nucleic acid sequence encoding such CBD protein, or hybridizes to nucleic acid sequence encoding such CBD protein and remains stably bound to it under stringent conditions.

The phrase "nucleic acid encoding the CBD protein of the present invention" includes nucleic acid of genomic, cDNA, synthetic, and semi-synthetic

origin which, by virtue of its origin or manipulation, is not associated with any portion of the polynucleotide to which it is associated in nature, and may be linked to a polynucleotide other than that to which it is linked in nature, and includes single or double stranded polymers of ribonucleotides, deoxyribonucleotides, nucleotide analogs, or combinations thereof, as long as the CBD being encoded retains its ability to bind cellulose with high affinity. The phrase also includes various modifications known in the art, including but not limited to radioactive and chemical labels, methylation, caps, internucleotide modifications such those with charged linkages (*e.g.*, phosphorothioates, phosphorodithioates, etc.) and uncharged linkages (*e.g.*, methyl phosphonates, phosphotriesters, phosphoramidites, carbamites, etc.), as well as those containing pendant moieties, intercalators, chelators, etc. as long as the CBD encoded by the nucleic acid retains the ability to bind cellulose with high affinity and in a reversible manner.

CBD encoding nucleic acid may be used to construct recombinant expression vectors capable of expressing the CBD protein or the CBD fusion protein of the present invention. A nucleic acid construct is capable of expressing a protein if it contains nucleotide sequences containing transcriptional and translational regulatory information and such sequences are "operably linked" to nucleotide coding sequences. "Operably linked" refers to a linkage in which the regulatory DNA sequences and the DNA sequence to be expressed are connected in such a way as to permit transcription and ultimately translation.

In constructing the CBD fusion protein expression vector, the nucleic acid encoding the CBD protein will

be linked or joined to the nucleic acid encoding the second protein such that the open reading frame of the CBD protein and the second protein is intact, allowing translation of the CBD fusion protein to occur. CBD nucleic acid may be obtained from a variety of cell sources that produce cellulose binding domains that bind with high affinity and in a reversible manner or that produce CBD encoding mRNA. The preferred source of CBD encoding nucleic acid is *Clostridium cellulovorans*. The CBD encoding nucleic acid may be obtained as described in Section 6.1.

The nucleic acid encoding the CBD protein of the present invention may be obtained from isolated and purified DNA from cell sources or by genomic cloning. Either cDNA or genomic libraries of clones may be prepared using techniques well known in the art and may be screened for particular CBD encoding nucleic acid with nucleotide probes that are substantially complementary to any portion of the gene. If detection of CBD protein encoding conserved nucleotide regions is desired, nucleotide probes should be based on CBD nucleotide sequences conserved from species to species. If detection of CBD protein encoding unique nucleotide regions is desired, nucleotide probes should be based on unique CBD nucleotide sequences. Alternatively, cDNA or genomic DNA may be used as templates for PCR cloning with suitable oligonucleotide primers. Full length clones, *i.e.*, those containing the entire coding region of the desired CBD protein may be selected for constructing expression vectors, or overlapping cDNAs can be ligated together to form a complete coding sequence. Alternatively, CBD-encoding DNAs may be synthesized in whole or in part by chemical synthesis using techniques deemed to be standard in the art.

Many vectors are available, and selection of the appropriate vector will depend on 1) whether it is to be used for nucleic acid amplification or for nucleic acid expression, 2) the size of the nucleic acid to be inserted into the vector, and 3) the host cell to be transformed with the vector. Each vector contains various components depending on its function (amplification of nucleic acid or expression of nucleic acid) and the host cell for which it is compatible.

The term "host cell" refers to those cells capable of growth in culture and capable of expressing a CBD protein or CBD fusion protein. The host cells of the present invention encompass cells in *in vitro* culture and include procaryotic, eucaryotic, and insect cells. A host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Expression from certain promoters can be elevated in the presence of certain inducers (*e.g.*, zinc and cadmium ions for metallothionine promoters. Therefore expression of the CBD protein or CBD fusion protein may be controlled. The ability to control expression will be important if the CBD protein or CBD fusion protein is lethal to a host cell. Modifications (*e.g.*, phosphorylation) and processing (*e.g.*, cleavage) of protein products are important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of protein. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the CBD protein or CBD fusion protein expressed. Preferably, the host

cell should secrete minimal amounts of proteolytic enzymes.

In the present invention, a host cell is provided comprised of nucleic acid encoding the CBD protein or CBD fusion protein of the present invention that is capable of binding to cellulose with high affinity. The preferred host cell for cloning and expression of the CBD proteins of the present invention is a prokaryotic cell. Procaryotes are particularly useful for rapid production of large amounts of nucleic acid, for production of single-stranded nucleic acid templates used for site-directed mutagenesis, for screening many mutants simultaneously, and for nucleic acid sequencing of the mutants generated. An example of a prokaryotic cell useful for cloning and expression of the CBD protein of the present invention is *E. coli* strain XL1-blue from Stratagene. Another example of a prokaryotic cell useful for cloning and expression of the CBD fusion protein is *Staphylococcus aureus*. Yet another example of a system useful for cloning and expression of the CBD protein or CBD fusion proteins is a *Pichia* expresion kit from Invitrogen Corporation (San Diego, CA).

Various expression vector/host systems may be utilized equally well by those skilled in the art for the recombinant expression of CBD proteins and CBD fusion proteins. Such systems include but are not limited to microorganisms such as bacteria transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing the desired CBD coding sequence; yeast transformed with recombinant yeast expression vectors containing the desired CBD coding sequence; insect cell systems infected with recombinant virus expression vectors

(*e.g.*, baculovirus) containing the desired CBD coding sequence; plant cell systems infected with recombinant virus expression vectors (*e.g.*, cauliflower mosaic virus CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (*e.g.*, Ti plasmid) containing the desired CBD coding sequence; or animal cell systems infected with recombinant virus expression vectors (*e.g.*, adenovirus, vaccinia virus) including cell lines engineered to contain multiple copies of the CBD nucleic acid either stably amplified (*e.g.*, CHO/dhfr, CHO/glutamine synthetase) or unstably amplified in double-minute chromosomes (*e.g.*, murine cell lines).

Vector components generally include, but are not limited to, one or more of the following: a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. The expression elements of these vectors vary in their strength and specificities. Depending on the host/vector system utilized, any one of a number of suitable transcription and translation elements may be used. For example, when cloning in procaryotic cell systems, promoters isolated from the genome of procaryotic cells, (*e.g.*, the bacterial tryptophane promoter) may be used. Promoters produced by recombinant DNA or synthetic techniques may also be used to provide for transcription of the inserted sequences.

A signal sequence may be a component of the vector, or it may be a part of the CBD nucleic acid that is inserted into the vector. The signal sequence may be the naturally occurring one preceding the CBD sequence or a non-naturally occurring sequence. The signal sequence should be one that is recognized and processed by the host cell. An origin of replication

refers to the unique site of initiation of replication of a host organism. It is desirable for cloning and expression vectors to comprise a selection gene, also termed a selectable marker. This gene encodes a protein necessary for the survival or growth of transformed host cells grown in a selective culture medium. Host cells not transformed with the vector containing the selection gene will not survive in the culture medium. Typical selection genes encode proteins that confer resistance to antibiotics or other toxins, *e.g.* ampicillin; complement auxotrophic deficiencies; or supply critical nutrients not available from complex media. One example of a selection scheme utilizes a drug to arrest growth of a host cell. Those cells that are successfully transformed with a heterologous gene express a protein conferring drug resistance and thus survive the selection regimen.

Expression and cloning vectors usually contain a promoter that is recognized by the host organism and is operably linked to the nucleic acid encoding the polypeptide of interest. Promoters are untranslated sequences located upstream (5') to the start codon of a structural gene (generally within about 100 to 1000 bp) that control the transcription and translation of a particular nucleic acid sequence, such as that encoding a CBD protein or CBD fusion protein, to which they are operably linked. Such promoters typically fall into two classes, inducible and constitutive. Inducible promoters are promoters that initiate increased levels of transcription from nucleic acid under their control in response to some change in culture conditions, *e.g.*, the presence or absence of a nutrient or a change in temperature. At this time, a large number of promoters recognized by a

variety of potential host cells are well known. These promoters are operably linked to nucleic acid encoding the polypeptide of interest by removing the promoter from the source nucleic acid by restriction enzyme digestion and inserting the isolated promoter sequence into the vector. Both the naturally occurring promoter sequence and many heterologous promoters may be used to direct amplification and/or expression of the polypeptide of interest. However, heterologous promoters are preferred, as they generally permit greater transcription and higher yields of expressed polypeptide of interest as compared to the naturally occurring promoter. In general, plasmid vectors containing promoters and control sequences which are derived from species compatible with the host cell are used with these hosts. The vector ordinarily carries a replication site as well as marker sequences which are capable of providing phenotypic selection in transformed cells. For example, *E. coli* is typically transformed using pBR322, a plasmid derived from an *E. coli* species (Bolivar, *et al.*, Gene 2:95 [1977]). The pBR322 plasmid contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid, must also contain or be modified to contain promoters and other control elements commonly used in recombinant DNA construction.

Promoters suitable for use with prokaryotic hosts illustratively include the β-lactamase and lactose promoter systems (Chang *et al.*, Nature, 275:615 [1978]; and Goeddel *et al.*, Nature 281:544 [1979]), alkaline phosphatase, the tryptophan (trp) promoter system (Goeddel Nucleic Acids Res. 8:4057 [1980] and EPO Appln. Publ. No. 36,776) and hybrid

promoters such as the tac promoter (H. de Boer *et al.*, Proc. Natl. Acad. Sci. USA 80:21-25 [1983]). However, other functional bacterial promoters are suitable. Their nucleotide sequences are generally known, thereby enabling a skilled worker operably to ligate them to nucleic acid encoding prorelaxin (Siebenlist *et al.*, Cell 20:269 [1980]) using linkers or adapters to supply any required restriction sites. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the nucleic acid encoding prorelaxin.

Expression vectors used in prokaryotic host cells will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA.

Construction of suitable vectors containing one or more of the above listed components and including the desired coding and control sequences employs standard ligation techniques. Isolated plasmids or nucleic acid fragments are cleaved, tailored, and religated in the form desired to generate the plasmids required.

Particularly useful in the practice of this invention are expression vectors that provide for the expression of prokaryotic cells of nucleic acid encoding the CBD protein. In general, expression involves the use of an expression vector that is able to replicate efficiently in a host cell, such that the host cell accumulates many copies of the expression vector and, in turn, synthesizes high levels of a desired polypeptide encoded by the expression vector.

Host cells are transfected and preferably transformed with the above-described expression or cloning vectors of this invention and cultured in conventional nutrient media modified as appropriate

for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences.

"Transformation" means introducing nucleic acid into an organism so that the nucleic acid is replicable, either as an extrachromosomal element or by chromosomal integration. Unless indicated otherwise, the method used herein for transformation of the host cells is the method of calcium treatment using calcium chloride as described by Cohen, F.N. *et al.*, Proc. Natl. Acad. Sci. (USA),69:2110 (1972).

For analysis to confirm correct sequences in plasmids constructed, the ligation mixtures are used to transform *E. coli* K12 strain 294 (ATCC 31446) and successful transformants selected by ampicillin or tetracycline resistance where appropriate. Plasmids from the transformants are prepared, analyzed by restriction and/or sequenced by the method of Messing *et al.*, Nucleic Acids Res. 9:309 (1981) or by the method of Maxam *et al.*, Methods in Enzymology 65:499 (1980).

Host cells may be transformed with the expression vectors of this invention and cultured in conventional nutrient media modified as is appropriate for inducing promoters, selecting transformants or amplifying genes. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

Prokaryotic cells used to produce the polypeptide of this invention are cultured in suitable media as described generally in Sambrook, *et al.* (1989) Electrophoresis buffers *in* Molecular Cloning (Nolan, C. ed.), Cold Spring Harbor Laboratory Press, NY, pp. B.23-24; Sambrook *et al.* (1989) Bacterial

Media *in* Molecular Cloning (Nolan, C. ed.), Cold Spring Harbor Laboratory Press, NY, pp. A.1-4.

The selection of host cells producing a CBD protein or CBD fusion protein of the present invention may be identified by at least four general approaches:

(a) DNA-DNA, DNA-RNA or RNA antisense RNA hybridization: the presence of nucleic acid coding for CBD proteins of the present invention can be detected by nucleic acid hybridization using hybridization probes and/or primers for PCR reactions comprising nucleotides that are homologous to the CBD coding sequence;

(b) the presence or absence of "marker" gene functions: the selection of host cells having nucleic acid encoding CBD protein of the present invention can be identified and selected based upon the presence or absence or certain marker gene functions, *e.g.*, resistance to antibiotics. For example, if the CBD coding sequence is inserted within a marker gene sequence of the cloning or expression vector, recombinants containing that coding sequence can be identified by the absence of the marker gene function. Alternatively, a marker gene can be placed in tandem with the CBD nucleic acid sequence under the control of the same or different promoter used to control the expression of the CBD coding sequence. Expression of the marker in response to induction or selection indicates expression of the CBD coding of the CBD coding sequence.

(c) assessing the level of transcription as measured by the expression of CBD protein or CBD fusion protein mRNA transcripts in the host cell: transcriptional activity of the CBD coding region can be assessed by hybridization assays. For example, polyadenylated RNA can be isolated and analyzed by

Northern blot using a probe homologous to the CBD coding sequence or particular portions thereof. Alternatively, total nucleic acids of the host cell may be extracted and assayed for hybridization to such probes; and

(d) detection of the CBD protein or CBD fusion protein as measured by immunoassay and, by the ability of the protein to bind cellulose with high affinity and in a reversible manner. The expression of CBD proteins can be assessed immunologically, for example by Western blots or by immunoassays such as RIAs. The expression of CBD protein can be assayed by the ability of the expressed protein to bind cellulose with high affinity and in a reversible manner.

The expressions "cell" and "cell culture" are used interchangeably and all such designations include progeny and ancestors. It is also understood that all progeny may not be precisely identical in DNA content, due to deliberate or inadvertent mutations. Mutant progeny that have the same function or biological activity as screened for in the cell are included.

The phrase "stringent conditions" as used herein refers to those hybridizing conditions that (1) employ low ionic strength and high temperature for washing, for example, 0.015 M NaCl/0.0015 M sodium citrate/0.1% SDS at 50 °C.; (2) employ during hybridization a denaturing agent such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl, 75 mM sodium citrate at 42 °C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M Sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 g/ml), 0.1% SDS, and 10% dextran

sulfate at 42 °C, with washes at 42 °C in 0.2 x SSC and 0.1% SDS.

The term "recovery" as used herein refers to the ability of the cellulose CBD protein complex to release the CBD protein from the cellulose-CBD protein binding complex under certain conditions, which conditions include the use of releasing agents, for example, denaturing reagents, such as 6M urea or 6M guanidine-HCl. Any releasing agent that has the ability of releasing the CBD protein from the cellulose-CBD protein binding complex can be used to recover the CBD protein. Preferably, the CBD is only temporarily denatured and not irreversibly degraded by treatment with the releasing agent. Thus, the CBD is recovered by reconstituting the eluted protein, as described in Section 7.1.1, 7.2.1, or 8.1.4.

The use of the phrase "cleaving agents" as used herein refers to a reagent used to cleave the CBD protein or CBD fusion protein specifically so as to release or excise certain components, such as the second protein of a CBD fusion protein, as desired. Suitable cleaving agents herein include enzymes, such as endoproteases, prohormone convertases, *e.g.*, PC1, PC2, furin, Kex2, subtilisin, or its mutants; and chemical agents, such as organic and inorganic acids, hydroxylamine, N-bromosuccinimide, and cyanogen bromide. Hydrolysis of peptide bonds catalyzed by a variety of proteolytic enzymes is taught in The Enzymes, 3rd Edition, Boyer, Ed., Academic Press, Vol. III [1971]; Meth. Enzymology., Vol. XIX, Perlman and Lorand, Ed., New York: Academic Press [1970]; Meth. Enzymol., Vol. XLV, Lorand, Ed. New York: Academic Press [1976]; Drapeau, J. Biol.Chem., 253:5899-5901 [1978] and Drapeau, Meth. Enzymology., 47:89-91[1977]. For an extensive listing of chemical agents, see

Witcop in Advances in Protein Chemistry, Anfinsen *et al.*, ed., 16:221-321, Academic Press, New York [1961], including Table III on p. 226. Other cleavage agents suitable herein are deemed to be understood by those skilled in the art keeping in mind the desired junction for cleavage and whether the reagent can act on reduced or oxidized forms of CBD fusion proteins. Conditions used for cleavage of the CBD fusion protein will depend on the cleavage agent employed, and the conditions will be readily apparent to one skilled in the art given the cleavage agent employed.

The CBD fusion protein of the present invention is designed and constructed to comprise the codon(s) necessary to achieve cleavage by the desired cleaving agent at desired positions, *i.e.* upstream, preferably adjacent the N-terminus of the second protein of the CBD fusion protein or downstream, and preferably adjacent the C-terminus of the second protein or both if the second protein of the fusion protein is an internal amino acid of the fusion protein.

The term "glycosylation" and grammatical derivatives as used herein refers to the post-translational modification process of adding a series of sugar residues to proteins to produce glycoproteins. Glycosylation can occur in the cytosol, the endoplasmic reticulum, or the Glogi apparatus of mammalian cells. Alternatively, glycosylation can be accomplished by synthetic methods, for example by providing an appropriate glycosyl donor. *See, e.g.*, Kahne, *et al.* J. Am. Chem. Soc., 111:6881-2 [1989].

This invention also relates to diagnostic detection of proteins of interest in test samples, especially in biological samples, such as tissue

extracts or biological fluids, such as serum or urine through use of the CBD fusion protein of the present invention. The biological samples are preferably of mammalian origin and most preferably of human origin. A preferred protein of interest to be detected in a mammalian biological sample is an HSP protein, an HSP antibody, cross-reactive HSP-related proteins, or antigenic portions thereof. The presence of the HSP antibody in a mammalian biological sample, for example, may be predictive or indicative of insulin-dependent diabetes mellitus (IDDM). In one embodiment of the present invention, the CBD Protein A fusion protein is comprised of a third protein, an IgG antibody, for example, IgG anti-HSP, which is used to detect the presence of an antigen, for example HSP, in biological samples using a variety of immunoassay formats well known in the art. Alternatively, the second protein of the CBD fusion protein is comprised of an antigenic determinant, an epitope, useful in the detection of antibodies that recognize the antigenic determinant. A preferred epitope is the HSP protein.

Protein A is a protein found in the cell wall of *Staphylococcus aureus* that binds to the Fc portion of IgG molecules and thus precipitates the antibodies. Protein A has utility in immunoassays, such as an RIA or ELISA, where it is used to isolate antibodies or antigen-antibody complexes.

In the present invention, Protein A is a preferred second protein of a CBD fusion protein. A CBD-Protein A fusion protein has utility in diagnostic immunoassays that detect the presence of or measure the quantity or concentration of an antibody or an antibody-antigen complex.

A CBD-Protein A fusion protein of the present invention also has utility in a diagnostic kit

comprised of cellulose and a CBD-fusion protein wherein the CBD fusion protein component retains its ability to bind both cellulose and IgG of a second component, for example, an antibody-antigen complex or an antibody. The CBD fusion protein of the present invention also has utility as a means for affinity purification of antibodies or antigenic determinants, *i.e.* epitopes. A preferred antigenic determinant of the present invention is the HSP protein, related protein or antigenic portion thereof. Preferred second proteins of a CBD fusion protein include HSP protein or anti-HSP IgG. In the present invention, CBD-HSP epitope fusion proteins find utility in immunoassays designed to measure quantities of HSP antibody found in the serum of human mammals.

In another embodiment of the present invention the CBD-fusion protein comprises a CBD component which retains its ability to bind cellulose or chitin and a second component which is an enzyme which retains its catalytic activity. The CBD-fusion proteins of this embodiment have utility as a means to permit enzyme catalyzed reactions on a solid substrate of cellulose or chitin. As described above, the CBD-enzyme fusion proteins can be prepared either by conjugating purified CBD to a purified enzyme preparation by standard chemical methods or by expression of CBD and enzyme encoding nucleic acids by recombinant methods. In one example of this embodiment of the invention, the second component of the CBD-fusion protein is a heparinase enzyme. The CBD-heparinase can be immobilized on a cellulose substrate and catalyzes conversion of heparin to component saccharide moieties.

Yet another aspect of the present invention relates to the ability of CBD to modify the growth of

plant tissues. Such ability finds application in the agricultural industry, for example, where it may be desirable to promote pollen tube growth and/or root growth. In addition, since CBD in high concentration has the ability to inhibit root growth it is useful to prevent the devlopment of undesirable plants, for example, "weed" plants. In another example, the CBD finds application in enhancing growth of fungi, such as mushrooms, wherein the CBD binds chitin and stimulates the growth and/or development of developing basidiocarps or fruiting bodies.

Yet another aspect of the present invention relates to the ability of CBD to bind to chitin of insect exoskeletons and other insect parts, including the mid-gut. Such ability finds application in the use of CBD as a bio-pesticide wherein the CBD is linked to a microorganism useful for control insect pests, such as *Bacillus thuringiensis* (BT) or other microorganism expressing the BT toxin gene, or such as a chitinase secreting fungus. The CBD functions to bind the microorganism to the chitinous portions of the insect where the microbially produced toxin or enzyme leads to death of the insect pest. In one mode of this embodiment of the invention, the CBD-microorganism complex can be applied directly to plant parts suspected or at risk of infestation by insect pests or, alternatively, the CBD-microorganism complex can be immobilized on bait stations which are set out in the infested fields or on plants infested with insect pests. In another alternative mode of this embodiment of the invention a CBD-microorganism complex is applied to wood or cellulose-based products to prevent or control infestation by wood destroying insect pests. For example, a CBD-microorganism complex is applied to telephone poles or wood or

cellulose-based construction components where the CBD component binds tightly to invading insects and the microorganism of the complex is effective to control the insect pests.

Still another aspect of the present invention relates to the ability of CBD to bind to chitin found in fungi, including fungi which are useful for bioremediation or degradation of toxic environmental pollutants. According to one mode of this embodiment, a CBD fusion protein is used in which the second component is an enzyme which degrades the environmental pollutant. The CBD-enzyme fusion protein can be immobilized or a cellulose substrate and the device used to remove the pollutant from the environment by enzymatic degradation, for example, by placing the device in a stream of surface or underground water. According to any alternative mode of this embodiment of the invention, the CBD is bound to a fungus which secretes an enzyme which is useful for biomediation *i.e.*, which degrades an environmental pollutant such as a pesticide. The CBD-fungus complex is then immobilized on a cellulose substrate and provides a useful device to enzymatically degrade an environmental pollutant such as a pesticide including DDT or such as wood containing products left over from logging or saw milling operations. In one illustrative example, CBD is complexed with a white rot fungus, for example *Phanerochaete chrysosporium*, which produces lignin peroxidase, manganese peroxidase and degrades the pesticide DDT. The CBD-fungus complex is useful for bioremediation, for example, to remove a pesticide such as DDT from the environment.

The "antibody" as used herein is meant to include polyclonal antibodies, monoclonal antibodies (MAbs), humanized or chimeric antibodies, single chain

antibodies, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. An epitope refers to an antigenic determinant of an antigenic molecule.

The term IgG refers to a class of antibodies. IgG is a tetramer containing two light chains and two heavy chains that represents 80% of all immunoglobulins.

The term "detectable label" as used herein refers to any label which provides directly or indirectly a detectable signal and includes, for example, enzymes, radiolabelled molecules, fluoresors, particles, chemiluminesors, enzyme substrates or co-factors, enzyme inhibitors, magnetic particles. Examples of enzymes useful as detectable labels in the present invention include alkaline phosphatase and horse radish peroxidase. A variety of methods are available for linking the detectable labels to proteins of interest and include for example the use of a bifunctional agent, such as 4,4'-difluoro-3,3'-dinitro-phenylsulfone, for attaching an enzyme, for example, horse radish peroxidase, to a protein of interest. The attached enzyme is then allowed to react with a substrate yielding a reaction product which is detectable.

Falling within the scope of the present invention is a signal amplification method wherein the use of a detectable label comprised of a CBD protein allows for detection of femtogram quantities of the substance of interest. In this method, a first CBD protein is part of a CBD fusion product that is incubated with a cellulose fiber under conditions suitable for formation of a cellulose-CBD fusion product binding complex. Excess labeled CBD, for example, CBD fused or bound to an enzyme, such as

horse radish peroxidase or alkaline phosphatase, or a fluoresor or chemical stain, is incubated with the cellulose-CBD fusion product binding complex under conditions suitable to allow binding of the excess labeled CBD. The binding of excess labeled CBD to the cellulose-CBD fusion product binding complex in effect allows for detection of very low quantities of substances of interest. In the signal amplification method of the present invention, the preferred cellulose fiber is a pebble-milled cellulose fiber. The pebble-milled cellulose fiber can be stained with a variety of chemical dyes or mixed with calcofluor that bind the cellulose and produce an intense bright blue fluorescence upon uv illumination.

Included within the scope of the present invention is a solid-phase polymerase chain reaction (PCR) wherein a first CBD protein is part of a CBD fusion product that is incubated with a cellulose fiber under conditions suitable for formation of a cellulose-CBD fusion product binding complex and wherein the thermocyclic amplification reactions take place on the binding complex.

Also falling within the scope of the present invention is the use of the Streptavidin/biotin detection system. Biotin is capable of forming a tight and, essentially, irreversible complex with Streptavidin. In this aspect of the present invention, a CBD fusion product comprised of a CBD protein fused to Streptavidin is provided. A nucleotide or protein is biotinylated through techniques deemed to be routine to those skilled in the art in order to form a biotinylated chimeric probe capable of binding a substance of interest. The biotinylated probe is incubated with a substance of interest, the CBD-Streptavidin is incubated with the

biotinylated probe and a detectable label is used to measure the biotinylated probe and consequently the substance of interest. The detectable label may be as described above in the signal amplification method, or any label, such as a fluoresor or chemical stain.

Also falling within the scope of the present invention is the therapeutic or diagnostic use of a CBD fusion product wherein the CBD fusion partner is a monoclonal antibody. For example, a CBD fusion product comprising (i) a CBD capable of binding to cellulose with high affinity and substantially free of other proteins with which it is naturally associated, and (ii) a monoclonal antibody capable of binding antigen can be used in a method to target a drug/cellulose complex or imaging agent/cellulose complex to a cancer cell producing the antigen. In this embodiment, a CBD-monoclonal antibody (or analogous antibody fragment, scFv fragment or antigen-binding region thereof) fusion product is administered to a mammal. Either concurrently with or following the administration of the CBD fusion product, a drug/cellulose or an imaging agent/cellulose complex is administered. Binding of the drug/cellulose or imaging agent/cellulose complex to the CBD-fusion product localized at the site of the antigen targets the drug or imaging agent to the relevant site for the desired therapeutic or diagnostic activity.

6. EXAMPLE: EXPERIMENTAL PROCEDURE FOR CLONING OF THE PUTATIVE CELLULOSE BINDING DOMAIN

6.1. Materials and Methods

6.1.1. Bacterial Strains and Plasmids

E. coli Xl 1 Blue strain was from StrateGene (La Jolla, CA), and was used for all cloning experiments. *E. coli* BL21 (DE3) and pET-8c were as

described (Studier *et al.*, J. Mol. Biol. 189:113-130 [1986]).

6.1.2. Materials

PC buffer, pH 7, contained 50 mM KH_2PO_4, 10 mM $Na_3C_6H_5O_7$ (sodium citrate), and 1 mM NaN_3. TEDG buffer (Chang *et al.*, J. Bacteriol. 172:3257-3263 [1990]) contained 10 mM Tris, pH 7, 0.1 mM EDTA, 0.1 mM dithiothreitol, and 5% v/v glycerol. Although Tris has a low buffer capacity at pH 7, the buffer was suitable because hydrogen ions were neither produced nor used. Restriction endonucleases were from BRL (Bethesda, MD). All other chemicals used were of the highest purity commercially available. Avicel® PH101 (lot #1117) was from FMC Corp. (Philadelphia, PA). Absorbant cotton was from the Seamless Rubber Co. (New Haven, CT). Cellulon® fiber was from Weyerhaeuser (Tacoma, WA). Granular chitin from crab shells is available from Sigma. All other binding substrates were purchased from Sigma Chemical Co., St. Louis, MO. Each of the polysaccharides was washed twice with PC buffer before use. Masses of the polysaccharides refer to their dry weights as supplied from the manufacturer, except for nigeran, Cellulon®, and cotton. These three binding substrates had large particle sizes which interfered with the assay. Nigeran was recrystallized by dissolving the solid in hot water, filtering, and cooling on ice. The sizes of the Cellulon® and cotton fibers were reduced by processing with a Gifford-Wood mini-mill for five minutes.

6.1.3. Cloning of Putative Cellulose Binding Domains

DNA primers complementary to the regions of *cbpA* (Shoseyov *et al.*, PNAS USA 89:3483-3487 [1992]) flanking the putative CBD (CbpA residues 28-189) were synthesized by a Gene Assembler Plus (Pharmacia). The forward primer contained an *Nco*I restriction site (recognition sequence: CCATGG) with the ATG in-frame with the gene fragment, to act as a translational start codon when cloned into the pET-8c vector cloning site. The reverse primer contained a stop codon and a *Bam*HI site. Polymerase Chain Reaction (PCR) was performed using 20 pmol each primer, 200 μM dNTPs, and 1 ng *cbpA* DNA [cloned into vector pGEMEX-1 (Promega) as in (Shoseyov *et al.*, 1992, *supra*)] as a template, in a total volume of 100 μl. Taq polymerase was from Amersham, using buffer conditions recommended by the manufacturer. PCR was carried out for 40 cycles as described (Innis *et al.*, Optimization of PCRs In: *PCR Protocols* Ed. Innis *et al.*, Pub. Academic Press, San Diego, pp. 3-12 [1990]). The PCR product was purified by phenol/chloroform extraction followed by ethanol precipitation and a wash with 70% ethanol, then dried by vacuum and resuspended in 27 μl distilled water. The DNA was then cleaved by *Nco*I and *Bam*HI, and run on a 2.5% low melting point agarose (Nuseieve GTG, FMC) gel in TBE buffer (Sambrook *et al.* (1989) Electrophoresis buffers *in* Molecular Cloning (Nolan, C. ed.), Cold Spring Harbor Laboratory Press, NY, pp. B.23-24. DNA bands stained by ethidium bromide were visualized by long-wave ultraviolet light and cut from the gel. The vector, plasmid pET-8c, was prepared by cleaving 1 μg of pET-8c DNA with *Nco*I/*Bam*HI and cutting the linearized DNA band from the gel. Vector and insert DNAs were ligated by using 100 ng of vector

DNA and 300 ng of insert with a Takara Ligation kit. The ligated plasmids were used to transform competent *E. coli* XL1-Blue strain, which were then plated on LB plates (Sambrook *et al.* (1989) Bacterial Media *in* Molecular Cloning (Nolan, C. ed.), Cold Spring Harbor Laboratory Press, NY, pp. A.1-4.) containing 100 μg/ml Ampicillin and 12.5 μg/ml tetracycline. After overnight incubation at 37 °C, colonies were selected and grown in liquid LB media with ampicillin and tetracycline. Plasmid DNA from each culture was rescued as described (Sambrook *et al.* (1989) *in* Molecular Cloning (Nolan, C. ed.), Cold Spring Harbor Laboratory Press, NY) and cleaved with restriction enzymes to verify the insertion of the gene fragment. The insert sequence was confirmed by DNA sequencing using the same procedures as those reported in Shoseyov *et al.* (1992) PNAS USA 89:3483-3487.

6.1.4. Expression of CBD

The overexpression vector (pET-CBD) enables us to overproduce the 17 kDa CBD in *E. coli* strain BL21(DE3). CBD was accumulated to at least 70 mg/liter in inclusion bodies. However, additional quantity of about 20 mg/liter of water-soluble CBD could be recovered from the water-soluble sonic extract of the *E. coli*. The cleared extract was mixed with (Sigmacell 20) cellulose; then the CBD-cellulose complex was washed by 1 M NaCl solution as well as distilled water to remove non-specific proteins, and then pure CBD was eluted by 6M guanidine-HCl. CBD was fully renatured by slow dialysis at room temperature and regained its ability to bind to cellulose (Figure 10. lane 2).

7. EXAMPLE: PURIFICATION OF CBD PROTEIN

7.1. Preparation of CBD Protein

7.1.1. Purification of CBD Protein

Plasmid DNA containing the insert was used to transform *E. coli* BL21 (DE3). Plasmid-containing cultures were grown at 37 °C in NZCYM (Sambrook *et al.* (1989) *in* Molecular Cloning (Nolan, C. ed.), Cold Spring Harbor Laboratory Press, NY) medium containing ampicillin (100 μg/ml) with shaking to Klett reading 160 (green filter). At this point, IPTG was added to a final concentration 1 mM. After 4 h, the cells were harvested by centrifugation, resuspended in 50 mM phosphate/12 mM citrate pH 7 (PC) buffer containing RNAse A at 10 μg/ml and DNAse I at 1 μg/ml, and lysed by sonication on ice with a Biosonic II sonicator at maximum power for 45 s followed by a 15 s cooling period, repeated a total of 4 times. The insoluble fraction of a 1 L cell culture was collected by centrifugation (30 min at 12,000 g, 4 °C) and resuspended in 20 ml of 6M guanidine HCl. This was kept on ice for 30 min with occasional vortexing to disperse the pellet. Insoluble debris was removed by centrifugation (30 min at 12,000 g, 4 °C). The soluble guanidine HCl extract was gradually diluted to 400 ml total volume with TEDG renaturation buffer over a two h period at 4 °C. Ammonium sulfate was added to 80% saturation. After four h at 4 °C, precipitated proteins were collected by centrifugation (30 min at 12,000 g, 4 °C), resuspended in 40 ml PC buffer, and dialyzed against PC buffer.

Further purification of the CBD protein fragment of CbpA was carried out by affinity chromatography on cellulose: Three additions of 1.0 g Avicel® PH101 microcrystalline cellulose were used to

remove the CBD protein from the solution. After each addition, the suspension was allowed to come to equilibrium (1 h at room temperature with slow rotation). The cellulose was then collected by centrifugation and removed before the next addition. The three grams of cellulose were washed once by 1 M NaCl/PC buffer and twice by PC buffer. Purified CBD was eluted from the cellulose by three washes with 10 ml 6M urea. The urea fractions were pooled and dialyzed against PC buffer (4 °C to about ambient temperature). Protein concentration in the final purified fraction was analyzed by colorimetric methods using the MicroBCA protein assay kit (Pierce, Rockford, IL), using bovine serum albumin (BSA) standards.

7.1.2. Determination of the CBD-Cellulose Dissociation Constant and the Cellulose Binding Capacity

Samples of CBD protein (typically 0 to 100 μg) were added to 1.5 ml capacity microfuge tubes containing PC buffer supplemented with 1 mg/ml BSA and the desired amount of cellulose (typically 1 mg added from a stock slurry containing 10 mg/ml cellulose and 1 mg/ml BSA in PC buffer). Potential competitors, *e.g.* cellobiose (4 mg/ml) or carboxymethylcellulose (CMC, 4 mg/ml) were included by adding 200 μl of a 20 mg/ml stock solution in PC/BSA buffer. The final volume was always 1 ml. The pH of the buffer was 7.0 unless otherwise noted. For experiments at other pH values, the PC/BSA buffer was adjusted by the addition of concentrated HCl or NaOH.

Assay tubes were mixed by slow vertical rotation (30 RPM) at 37 °C for one h. The samples were then spun in a microfuge for one min to sediment the cellulose and cellulose-protein complexes. After

removing the buffer, the pellet was washed by resuspension in 1 ml of PC buffer. The wash was separated out by centrifugation and discarded. Pellets were then resuspended in a final 1 ml PC buffer. (The centrifugation step would not be expected to perturb the equilibrium because the [C] and [PC] were concentrated to the same extent.)

Of the original BSA in the assay tubes (ca. 1 mg/ml), only about 0.1 μg would remain after the washing steps assuming no non-specific adsorption and a liquid volume of 10 μl in the pellet. Any color development due to this residual BSA was accounted for by the 0 CBD control tubes. Aliquots (150 μl) of this well-mixed suspension were taken for protein determinations with the MicroBCA kit. The manufacturer's instructions were followed, except that the sample volume was brought to 0.5 ml with PC buffer, to which 0.5 ml of BCA working reagent was added. Assay mixtures were incubated at 60 °C for 30 min. The protein concentration was determined colorimetrically from the cleared supernatants at 562 nm in a Shimadzu 160 U spectrophotometer. Assay tubes to which no CBD protein was added were used to correct for a small amount of color development caused by the cellulose and residual BSA. The data were compared to BSA standards and adjusted to accommodate the dilutions that were made to determine the amount of protein bound to the cellulose in each sample. The practical detection limit of this assay was about 0.2 μg/ml. After correction for dilutions, this corresponds to about 0.034 nmol of CBD bound to the cellulose in the assay tube. The free CBD protein concentration, [P], was determined by subtracting the bound protein concentration, [PC], from the total CBD added to the tube, $[P]_t$:

$$[P] = [P]_t - [PC] \quad (1)$$

The system was analyzed assuming a simple equilibrium interaction (Segel, 1975):

$$P + C \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} PC \quad (2)$$

where the dissociation constant, K_d, is defined as:

$$K_d = \frac{k_{-1}}{k_1} = \frac{[P]\,[C]}{[PC]} \quad (3)$$

The data were analyzed by double reciprocal plots of 1/[PC] versus 1/[P] at different fixed levels of cellulose (equation 4) and W:

$$\frac{1}{[PC]} = \frac{K_d}{[PC]_{max}} \cdot \frac{1}{[P]} + \frac{1}{[PC]_{max}} \quad (4)$$

and by Scatchard plots of [PC]/[P] versus [PC]:

$$\frac{[PC]}{[P]} = \frac{-1}{K_d}[PC] + \frac{[PC]_{max}}{K_d} \quad (5)$$

It must be noted that the cellulose is not a soluble component, and that [C] represents the concentration of binding sites on the cellulose surface exposed to the buffer, per unit volume. Similarly, the [PC] represents the concentration of binding site-protein complexes per unit volume. Straight lines were fitted to the data points by the

least-squares method using DeltaGraphProfessional plotting application (Deltapoint, Inc., Monterey, CA). Each point was the average of three independent protein assays from the same binding assay tube. Experiments were performed in duplicate. At least two different amounts of cellulose were used to determine the K_d and PC_{max}/g cellulose. These values were averaged to provide the listed values in Table I, Figure 9.

Experiments were performed to determine the effect of temperature on CBD binding to cellulose. 200 µl of CBD (40 µg) were added to 300 µl of 50 mM PC buffer pH 6, 1 M NaCl (final concentrations) and 1 mg of cellulose. Several mixtures of the CBD/cellulose mixtures were incubated at different temperatures (16 to 60 °C) for 1 hour with occasional mixing and were then centrifuged for 5 min at 10,000 g. The pellets were washed twice with PC buffer pH 6 containing 1 M NaCl. 40 µl of 6 M guanidine-HCl were added to the pellet that was suspended and further incubated for 30 minutes at room temperature. After centrifugation, 20 µl samples were diluted in 780 µl ddH_2O and protein concentration was determined using a Bio Rad kit.

7.1.3. Determination of Binding to Other Polysaccharides

Xylan, nigeran, Sephadex® G-75, and chitin were used in assays to determine whether they were substrates for CBD protein. In all cases, the methods used were the same as those used in determining the binding to cellulose. Chitin exhibited a very high background in the MicroBCA assay, which increased proportionally to the 60 °C incubation time, so the color development time was reduced to 15 min. Because of chitin's high background, only two widely different protein concentrations were used.

7.2. Results

7.2.1. Purification of the CBD for Binding Analyses

In order to selectively produce the putative CBD region of CbpA (residues 28-189), oligonucleotide primers were designed complementary to bases 67 to 86 and 558 to 579 of *cbpA* (Figure 1). As shown in Figure 2, these primers were designed with mismatches to create an *Nco*I site and an ATG start codon on one end of the PCR product and a TAG stop codon followed by a *Bam*HI site at the other end. This gene fragment was then cloned into the T7 RNA polymerase expression plasmid pET-8c, resulting in plasmid pET-CBD. *See*, Studier, F., and B.A. Moffatt (1986) J. Mol. Biol. 189:113-130. The cloned gene fragment codes for a methionine at the N-terminus of the CBD, but the rest of the CBD aa sequence corresponds to residues 28 to 189 of CbpA. The protein fragment has a molecular weight of 17634. The insertion was verified by DNA sequencing. CBD protein was produced by *E. coli* BL21 (DE3) cells harboring pET-CBD. After the addition of IPTG, this host strain produces T7 RNA polymerase, which recognizes the T7 promotor in the pET vector. The *cbd* gene fragment was under the control of this inducible promotor, and CBD protein was synthesized in large amounts after induction (Figure 3). After a four h production period, the soluble extract from the lysed cells contained only small amounts of CBD protein, while most was found in the insoluble fraction. This protein was readily soluble in concentrated guanidine hydrochloride, and was renatured by slow dilution into TEDG buffer. It was found that protein prepared in this fashion binds to Avicel®, verifying the putative CBD. Although this fraction is mostly CBD protein, the assays described

require the protein to be highly pure. This purity is provided by a single cellulose-affinity binding step, as described in the Section 7.1.1. The affinity-purified CBD protein appears on acrylamide gels as a single band when stained with Coomasie brilliant blue. Approximately 70 mg of CBD protein can be recovered from the cells harvested from a 1 L culture.

7.2.2. Time course of the Binding of the CBD to Cellulose and Effect of Temperature

The time course of the interaction of Avicel® with CBD (Figure 4) discloses several features of the process:

(a) At initial concentrations of 1.0 mg/ml Avicel® and 2.0 μM CBD (*i.e.*, $[P]_o$), a plateau value of 1.2 μM complex (*i.e.*, [PC]) is attained by 60 minutes. A separate experiment established that the maximum CBD binding capacity of the cellulose sample was 2.1 μmoles x g^{-1}, corresponding to an effective concentration of 2.1 μM total cellulose sites (*i.e.*, $[C]_0$). Assuming that an equilibrium was established (verified below), K_d defined as [P][C]/[PC] is about 0.6 μM.

(b) The second-order rate constant for association (k_1) calculated from the integrated rate equation for a reversible Bi Uni reaction (Capellos & Bielski, 1980; Wilkinson, 1980) is about 2.7 x 10^4 M^{-1} x min^{-1} (average value for points from 5-60 min). The rate constant for the dissociation of the complex (k_{-1}) calculated as $k'K_d$ was 1.6 x 10^{-2} x min^{-1} ($t_{1/2}$ = 43 min). The relatively long $t_{1/2}$ for complex dissociation permitted the C+PC pellet to be washed once without significant loss of bound CBD. (Resuspension and recentrifugation of the initial pellet was completed

in less than 1 min. During this period, less than 3% of the bound CBD would be lost.) It was also observed that after prolonged incubation, the measured [PC] declined, dropping to a level of about 50% of the maximum value after 18 hours. This decline may be caused by gradual denaturation of the protein. To reduce artifacts resulting from these effects, we used the shortest incubation time for which equilibration appeared to be "complete". (Any further increase in binding beyond 60 min would be obscured by the experimental error.) CBD binding to cellulose was 4.5 +/-0.5 mg/gr cellulose and was not affected by a temperature range of about 16 to about 60 °C.

7.2.3. Analysis of the CBD-Cellulose Binding Affinity and Binding Capacity

Figure 5 shows a typical diagnostic plot of the binding of pure CBD to Avicel® microcrystalline cellulose. Within experimental error, the plots were linear yielding a K_d of about 0.6 μM and a $[PC]_{max}$ of 2.1 μmoles CBD bound per gram Avicel®. The latter value corresponds to approximately 37 mg of CBD protein per gram of Avicel®. The linearity of the diagnostic plots suggests that only one type of CBD-cellulose interaction is occurring.

The ability of CBD to bind cellulose types other than Avicel® was also investigated. Table 1, Figure 9, shows the values for the K_d and PC_{max} found for each of the substrates. Sigmacell 20 and 50 are described as microcrystalline forms of cellulose; these materials were also found to bind to CBD. Highly crystalline forms of cellulose such as absorbant cotton and Cellulon® fiber (crystalline cellulose from *Acetobacter xylinum*) were able to bind substantially more of the CBD (up to 6.4 μmol CBD per

gram of substrate). Fibrous and microgranular cellulose, however, which are more processed and thus contain less of the native crystalline form, were found to bind a smaller amount of the CBD. The CBD-cellulose dissociation constant was about the same for all forms of cellulose, while the PC_{max} varied over a 30-fold range.

7.2.4. Binding Site Competition

To determine if soluble carbohydrates competed with Avicel® for the CBD protein, cellobiose (a β-1,4 linked glucose dimer) and CMC (a soluble derivative of cellulose) were included in some assays at 4 times the weight/volume of Avicel® (1 mg Avicel®, 4 mg cellobiose or CMC per ml assay). As Table 1, Figure 9, shows, no significant differences in the K_d or PC_{max} were observed, indicating these soluble carbohydrates did not affect the binding of the CBD to Avicel®.

7.2.5. Effect of pH on the Dissociation Constant

C. cellulovorans is a neutrophilic organism, thriving only around pH 7 (Sleat *et al.*, 1984), so this pH was used for most of the binding assays. However, other experiments established that the K_d and PC_{max} did not vary significantly with changes in pH over the range 5.0 to 8.0. In addition, it was noted that PC buffers as acidic as pH 3.5 or as basic as 9.5 would not remove the CBD from Avicel® during 1-min washes.

7.2.6. Binding of the CBD to Oth r Polysaccharides

Xylan, Sephadex® G-75, nigeran, and chitin were used to explore the binding specificity of the CBD. Of these, only chitin showed measurable binding of the CBD peptide Table 1, Figure 9. The chitin-CBD K_d and binding capacity were similar to the Avicel®-CBD values.

7.3. Results

Our results show that CbpA contains a domain responsible for cellulose binding, and that this domain could function independently from the rest of CbpA. Because the purification protocol employed denaturation and renaturation steps, the fact that the purified protein was functional indicates that the CBD protein sequence was sufficient for proper folding of the protein fragment.

We have found non-specific binding of the CBD to the assay tubes to be a problem in performing equilibrium binding experiments, and have developed an assay in which the CBD and cellulose are equilibrated in the presence of excess BSA. The BSA effectively eliminates non-specific CBD interactions with the tube. After equilibrium is reached, the cellulose and cellulose-protein complexes are collected and washed, then assayed for bound proteins. As described earlier, the dissociation of the CBD-cellulose is slow so that no detectable amount is removed during a rapid wash step.

The bound CBD was measured directly by the protein assay, and the free CBD was calculated by subtracting the bound CBD from the total CBD, as shown in equation 1, Section 7.1.2. This method has the

advantage that any CBD molecules adsorbed *non-specifically* with low affinity to the cellulose would be removed by the wash step, resulting in data that more accurately reflect the specific, high affinity interaction between the CBD and the cellulose surface. As shown in Figure 5, data gathered using this type of assay yields (within experimental error) linear diagnostic plots. Double reciprocal plots are a convenient and conventional way of determining binding affinities and capacities for reversible Bi uni systems. The validity of the assay is supported by the observation that PC_{max} increases linearly with the amount of cellulose used, while K_d is independent of the cellulose quantity. Table 1, Figure 9, shows the results obtained with several forms of cellulose and other carbohydrates. The results indicate that cellulose types described as "crystalline" have a higher CBD binding capacity than highly processed celluloses that have lost much of their crystallinity. The fact that the PC_{max} of cellulose samples vary widely with different cellulose types while the K_d remains constant indicates that we have measured one type of strong protein-cellulose interaction occurring between the CBD and the cellulose. The lower PC_{max} of highly processed celluloses reflects a smaller number of potential protein interaction sites in the sample, and seems to correlate with the crystallinity of the sample. This result would indicate, although not wishing to be limited by theory, that there is some special feature present in crystalline cellulose that makes it acceptable as a binding substrate, whereas amorphous cellulose is found lacking.

To further characterize the substrate specificity of the CBD, we measured the effect of added soluble substrates (cellobiose or CMC) on

cellulose binding. Excess cellobiose or CMC had no effect on the CBD-Avicel® K_d or PC_{max}, as shown in Table 1, Figure 9. This lack of competition suggests that the CBD recognition site is specific for something more complex than a simple repeating glucose or cellobiose moiety, and further suggests that, perhaps, a particular three-dimensional arrangement of cellulose chains is needed.

The specificity of the CBD for crystalline cellulose prompts a consideration of chitinases, which are known to bind tightly to chitin, a polymer of N-acetylglucosamine in β-1,4 linkage. Like cellulose, chitin comes in a variety of forms, depending on the source and purification method used in its isolation (Cabib (1988) Methods Enzymol. 161:460-462; Blackwell (1988) Methods Enzymol. 161:435-442). The chitin used for affinity purification of chitinases is α-chitin, in which the chains are arranged in an antiparallel fashion. This form of chitin is crystalline with a structure similar to that of native crystalline cellulose (often referred to as cellulose I). Cellulose I is the form in which the cellulose chains are arranged in parallel bundles, as opposed to cellulose II, in which the chains are in an antiparallel configuration. Processing of cellulose I under harsh conditions causes its disruption resulting in cellulose II. Both forms are crystalline, due to extensive hydrogen bond formation. Since our isolated CBD binds to less processed forms of cellulose, *i.e.*, largely cellulose I, we were interested to see if the CBD would bind to α-chitin, which has a similar crystal structure, although of opposite strand orientation. Surprisingly, we found that the CBD did accept chitin as a binding substrate with a K_d very similar to that for cellulose. Xylan (β-1,4 xylose),

nigeran (alternating α-1,4 and α-1,3 glucose), and Sephadex® G-75 (α-1,6 glucose with α-1,3 branches) (Coutinho *et al.* (1992) Mol. Microbiol. 6:1243-1252) were also examined, but the CBD did not show measurable binding to any of them under the conditions of the assay. Since chitin is the only one of these substrates that is crystalline, we feel that this demonstrates the importance of crystallinity in the substrate.

Although this is certainly a part of the benefit of tight substrate binding, recent studies (Din *et al.* (1991) Bio/Technol. 9:1096-1099) have shown that binding to cellulose by isolated, non-enzymatic CBDs causes disruption of the cellulose fibers in a non-hydrolytic fashion. It is thought that the protein-cellulose binding lowers the degree of inter-chain hydrogen bonding near the surface of the crystal, which is followed by extensive hydration of these cellulose chains. The end result is a decrease in crystallinity. This process, termed "amorphogenesis," renders more of the cellulose fiber accessible to the endo-β-1,4-glucanases. As we have discovered, the CBD exhibits no such amorphogenic effect.

7.4. CBD Binding to *Aspergillus niger*, *Spodoptera littoralis*, and *Heliothis armigera*

Aspergillus niger mycelia (approximately 20 mg) were incubated for 1 hour at room temperature with 400 μl CBD (120 μg) in 20 mM Tris-HCl pH 7. The mycelia were washed twice with the same buffer and then the pellet was resuspended with SDS-PAGE sample application buffer containing ß-mercaptoethanol and SDS. The mixture was boiled for 5 min and the

proteins were analyzed by SDS-PAGE. Figure 16 shows that CBD binds to the mycelia of *Aspergillus niger*.

Mid-gut membranes of *Spodoptera littoralis* and *Heliothis armigera* (major insect pathogens in agriculture) were isolated under a microscope. Approximately 20 mg were incubated for 1 hour at room temperature with 400 µl CBD (120 µg) in 20 mM Tris-HCl pH 7. The chitin containing membranes were washed once with 1 M NaCl in the same buffer then washed twice with the same buffer without NaCl and then the pellets were resuspended with SDS-PAGE sample application buffer containing B-mercaptoethanol and SDS. The mixtures were boiled for 4 minutes and the proteins were analyzed by SDS-PAGE. The results are shown in Figures 17 and 18. The results indicate that the CBD binds to the mid-gut membrane of *Spodoptera littoralis* and *Heliothis armigera*.

8. CONSTRUCTION OF CBD-ProtA GENE AND ITS EXPRESSION IN *E. COLI*

8.1. Material and Methods

8.1.1. Enzymes and Chemicals

Chemicals were purchased from Sigma Chemicals Inc. unless stated otherwise. Restriction enzymes were purchased from New England Biolabs, Inc. Taq polymerase was purchased from Promega, Inc.

8.1.2. Plasmids and Bacteria

The plasmid pCB1 carrying *cbpA* (Shoseyov *et al*. 1992) was used to amplify *cbd* by PCR methods. The expression vector pRIT2 (Nilsson, et al (1985) EMBO J. 4(4):1075-1080) was used for the construction of the fusion genes. Initial transformations were conducted using *E. coli* strain XL1-blue (Strategene). Expression of CBD-ProtA was conducted in the

temperature sensitive repressor containing strain, 2097. Also, *E. coli* strain N4830-1 carries the temperature-sensitive repressor ci857 and is equivalent to *E. coli* strain 2097. The *E. coli* strain N4830-1 is available from Pharmacia, Inc.

8.1.3. Cloning of CBD-ProtA

CBD was PCR amplified using *cbpA* gene as a template. Primers A (N-terminal primer; '5-GGGGGAATTCCATGGCAGCGACAT-'3) contain EcoRI site, primer B (C-terminal primer; '5-GGGGGGATCCTATGGTGCT-'3) contain a stop codon followed by a BamHI site. The primers were designed and synthesized in a way that enable force cloning of EcoRI/BamHI 500 bp DNA fragment into the plasmid pRIT2 "in frame" fused to the C-terminal part of Protein A gene. PCR conditions were as described by Innis and Gelfand (1990) with the following modifications: 2 ng template DNA (*cbpA*) and 1mM $MgCl_2$ were used in the reaction mixture. The reaction was conducted using a programmable thermal controller (M&J Research Inc.). Standard DNA manipulations were conducted according to Sambrook et al (1989) *in* Molecular Cloning (Nolan, C. ed.), Cold Spring Harbor Laboratory Press, NY. The PCR amplified product was digested with EcoRI and BamHI, and the expected 500 bp DNA fragment (Figure 1) was isolated from 1.5% agarose gel using QIAEX gel extraction kit (Qiagen Inc.). The EcoRI/BamHI fragment was ligated into EcoRI/BamHI linearized pRIT2 using T4 Ligase. The ligation mixture was used to transform XL1-Blue competent cells and transformed colonies were selected on LB agar plates containing 100 mg/L ampicillin. Successful construct containing the DNA insert was designated pCBD-ProtA1.

8.1.4. Expr ssion and purification of CBD-ProtA

Expression of the fusion protein was conducted as described elsewhere (Nilsson, et al (1985) EMBO J. 4(4):1075-1080) .40 ml LB containing 50 mg/L ampicillin inoculated with 400 ul of overnight culture of *E. coli* strain 2097 containing pCBD-ProtA1. The culture was grown to 100 Klett (green filter) at 30 °C and then shifted to 42 °C for 45 minutes and then grown for additional 2 hours at 40 °C. The cells were harvested by centrifugation at 2,000 g for 10 min. and resuspended in 10 ml of 20 mM Tris-HCl buffer pH 7. The cells were lysed using W-385 sonicator (Heat Systems-Ultrasonic, Inc.) at maximum power for 1 minute followed by a 30-second cooling period, repeated 3 times. The lysate was cleared by centrifugation (4,000 g for 10 min) and 500 mg of cellulose particles (Sigmacell 20, 20 microns average particle size) were added. The suspension was incubated for ten minutes at RT and centrifuged (2,000 g for 10 minutes). The supernatant was removed and the pellet was washed once with 5 ml of 1 M NaCl to remove non-specific binding proteins and twice with 10 ml of deionized water. CBD-ProtA was remove from the cellulose by 5 ml 6M guanidine-HCl. After centrifugation (2,000g, 10 min) the solution was then dialysed against 20 mM Tris-HCl buffer (pH 7). Proteins were analyzed on 12.5% SDS-PAGE according to Laemmli (1970).

8.1.5. Binding Analysis of CBD-ProtA

The binding of CBD-ProtA to IgG was determined as follows: A 100 μl suspension of rabbit IgG (H+L)-Sepharose (8 mg/ml, Bio-Makor Inc.) that was

prewashed with PBS (Phosphate Buffer Saline pH 7.4) was mixed with 1 ml of isolated CBD-ProtA (50 μg/ml). The mixture was incubated for 1 hour at 8 °C and then centrifuged (2,000g, 5 minutes). The supernatant was removed and the pellet was washed twice with 500 μl of PBS. Then, the CBD-ProtA was eluted with 200 μl of 1M acetic acid. The pellet was mixed with 15 μl of sample application buffer (SAB; 125 mM Tris-HCl pH 6.8, 4% SDS, 20% glycerol and 0.002% bromophenol blue), 15 μl samples of the different fractions were mixed with 15 μl of SAB, then boiled for 5 minutes and analyzed on SDS-PAGE.

The binding of CBD-ProtA to cellulose was determined as follows: 20 mg of cellulose (Sigmacell 20) were mixed with 200 μl of isolated CBD-ProtA (50 μg/ml). The mixture was incubated for 15 min at RT and centrifuged at 2,000g for 5 min. The pellet was washed with 200 μl 1 M acetic acid and the pellet was resuspended with 40 μl of SAB. The cellulose suspension was boiled along with the 15 μl of the acetic acid wash (mixed with 15 μl SAB) and analyzed on SDS-PAGE.

8.1.6. Results

It was shown that CBD-ProtA can be expressed in *E. coli*, and purified using cellulose in a one-step purification. The fusion protein has the expected size, 45 kDa (Figure 8). CBD-protA retained its affinity to cellulose as well as to IgG. It was shown that 1 M acetic acid releases the CBD-ProtA:IgG bond but not the CBD-ProtA:cellulose bond.

Using this expression vector, we produced 6 mg of CBD-ProtA per 1 liter of culture. Our experience with the T7 overexpression system enabled

us to produce more then 10 times that amount (70 mg/l) of pure CBD (Figure 10).

9. CLONING OF CBD-HSP FUSION PROTEIN

An example for the cloning of CBD-HSP fusion protein: PCR primers for the amplification of HSP gene are prepared using the plasmide SJ60 as a template. The vector was described by Jindal *et al.* (1989) Mol. Cell. Biol. 9:2279. The primers will contain KpnI site at the N terminal of HSP and stop codon followed by a BamHI site at the C terminal.

Forward primer: 5'-ACGGTACCACTTCGGTTACCCACAGTC-3'

Reverse primer: 5'-GGGGATCCTACATGCCACCTCCCATTAG-3'

In order to enable translational fusion of the C terminal part of CBD to the N terminal part of HSP, we introduce a Kpnl site at the 3' end of cbd gene. This introduction achieves PCR amplification of cbd using pET-CBD as a template and the following primers:

Forward primer: 5'-GTATACCAGCCATGGCAGCG-3'

Reverse primer: 5'-GTACATCTGGATCCTATGGTACCGT-3'

The amplified DNA is digested with NcoI and BamHI and is then ligated into NcoI/BamHI predigested pET8c vector. The ligation mixture is then used to transform XL1Blue and the new plasmid designated as pET-CBDK. This plasmid is digested with Kpnl and BamHI and the KpnI/BamHI restricted HSP-PCR amplified fragment is ligated; transformed into XL1Blue and after confirmation of the construct will be used to transform BL21(DE3) for overproduction of CBD-HSP.

10. CONSTRUCTION OF NH_2-V_H-V_L-CBD-CO_2

CBD fused to recombinant antibody is carried out by cloning any desired V_H-V_L using the "Recombinant Phase Antibody System" (Pharmacia Inc.). The resulting pCANTAB5 plasmid carrying the V_H-V_L is used as a template for PCR amplification using the following primers:

Forward primer: 5'-AGCCATGGCGGCCCAGC-3'

Reverse primer: 5'-GGGGTACCAACAGTTTGTGCGGCC-3'

These primers introduce the NcoI site at the 5' of the V_H-V_L and KpnI site at the 3' end. The amplified fragment is digested (partially if necessary) with NcoI and KpnI and is used in the expression vector of C-terminal fusion of CBD, below.

To enable translational fusion of the N terminal part of CBD to the C terminal part of V_H-V_L, we introduce a KpnI site at the 3' end of cbd gene. This introduction is achieved by PCR amplification of cbd using pET-CBD as a template and the following primers:

Forward primer: 5'-GGGCCATGGCAGGTACCTCATCA-3'

Reverse primer: 5'-GTACATCTGGATCCTATGGTGCTGT-3'

These primers introduce KpnI sit at the 5' end of cbd gene after the NcoI site, and maintain the stop codon followed by BamHI site at the 3' end. The amplified DNA is digested with NcoI and BamHI and then is ligated into NcoI/BamHI predigested pETSc vector. The ligation mixture is used to transform XL1Blue and the new plasmid is designated as pET-KCBD. This plasmid is digested with NcoI and KpnI and the NcoI/KpnI restricted V_H-V_L amplified fragment is then ligated; transformed into SL1Blue and after conformation of the construct is used to transform

BL21(DE3) for overproduction of V_H-V_L-CBD fusion protein.

In view of the above-disclosure and what is generally known in the art, it would be apparent to one of ordinary skill that a wide variety of CBD fusion products can be prepared which comprise the CBD and second proteins of known sequence.

11. IgG PURIFICATION BY CBD-ProtA CELLULOSE

This example demonstrates the use of CBD for purification of IgG. Human serum (500 µl, 1:10 diluted in Phosphate Buffered Saline (PBS), pH 7.4 was added to 20 mg suspension of CBD-ProtA-Cellulose (4 µg/mg). The mixture was incubated for 1 hour at 8 °C and then the unbound protein was removed by a quick centrifugation and two 500 µl PBS washes. IgG was recovered by 200 µl 1 M acetic acid wash (15 minutes at room temperature) while CBD-ProtA remained bound to the cellulose.

12. CHARACTERIZATION OF CBD

This example demonstrates that CBD does not disrupt cellulose. CBD has no cellulose disruption/amorphogenesis activities. Approximately 1 mg of CBD was incubated overnight at 37 °C with 10 mg portions of two different types of cellulose (cotton and Cellulon®) in 20 mM Tris-HCl buffer pH 7. The cellulose samples were gold-coated and observed under scanning electron microscope. Figures 14A and 14B demonstrate clearly that CBD-treated cellulose fibers remained intact compared to the control, untreated cotton.

The effect of pH and NaCl concentrations on the binding capacity of CBD-ProtA to cellulose was examined. 450 µl of CBD-ProtA (100 µg) were added to

a tube containing 550 μl of different buffers (50 mM citrate phosphate containing different NaCl concentrations *i.e.*, 0, 0.25, 0.5, 1 M NaCl, 1 mg of cellulose (Sigmacel 20), and 1 mg BSA. The tubes were incubated for 1 hour at 37 °C with occasional mixing and were centrifuged for 5 minutes at 10,000 g. The pellets were washed twice with their corresponding buffers. 40 μl of 6 M guanidine-HCl were added to the pellet that was suspended and further incubated for 30 minutes at room temperature. After centrifugation, 20 μl samples were diluted in 780 μl ddH_2O and protein concentration was determined by a BioRad kit. Figure 15 shows that at low pH values, increasing concentrations of NaCl increased the binding of CBD-ProtA to cellulose up to 70 mg/gr cellulose.

13. CBD MODIFIES THE GROWTH OF PLANT TISSUES

13.1. Material and Methods

13.1.1. Plant Material

Flowers of peach (*Prunus persica* cv. Texas) were obtained from a plot near Rehovot. Anthers collected from flowers at the first day of anthesis were excised from the filaments and dehydrated at 30 °C for at least 24 hours. The pollen released was used fresh or stored at -20 °C. Seeds of *Arabidopsis thaliana* were obtained from the University stock.

13.1.2. Pollen Germination *In Vitro*

Pollen grains were germinated in liquid cultures of 100 μl growth medium in each Eppendorf tube. The growth media were made as follows: Peach pollen grains were germinated in growth medium containing 15% (w/v) sucrose, 100 μg/ml H_3BO_3, 200 12 μg/ml $MgSO_4.7H_2O$ and 200 μg/ml $Ca(NO_3)_2.4H2O$. Different concentrations of CBD or BSA were added to the tubes.

The amounts of pollen grains in each tube were approximately 1000. For each treatment, three repetitions were made. After incubation overnight at 25 °C in a dark chamber, the pollen was examined in populations of a least 100 pollen grains in each specimen (300 per treatment).

13.1.3. Seed Germination

Seeds of *Arabidopsis taliana* were washed in distilled water. About 100 seeds per treatment were soaked in 1 ml of distilled water in glass culture tubes of 2 cm diameter and 10 cm long. The tubes were placed in a growth chamber in which the temperature was constant at 25 °C and photoperiod of 16/8 hours of light/dark respectively. After four days, the length of the seedlings, shoots, roots and root hairs (the longest in each seedling) were measured to examine the effect of the different treatments on the plant organs.

13.1.4. Histochemical Observation

Peach pollen tubes that were grown with or without CBD were washed from the growth media by pelleting for 1 minute at 10,000g fixed overnight using 4% glutaraldehyde in 0.1M phosphate buffer pH 7.4. The pollen tubes were washed with the same buffer, then carefully resuspended in the same volume of 4% glutaraldehyde in the same buffer and kept overnight at 4 °C. The pollen tubes were pelleted again, washed with the same buffer, with distilled water and then stained with white fluorescent brightener (calcofluor, 0.1% in 0.1M K_3PO_3), to indicate crystalline cell wall components. Microscopic observations were made by fluorescent light microscope (RA Zeiss Oberkochen). *Arabidopsis*

seedlings were treated in the same way, however the solutions were changed without pelleting.

Peach pollen tubes and *Arabidopsis* seedlings were examined by immunogold-silver stain (IGGSS) to indicate CBD attachment to cellulose. The plant material was first fixed overnight 4% glutaraldehyde in PBST (Phosphate buffered saline TWEEN 2™). The specimens were washed with PBST for one hour, soaked for one hour in 1% skim milk, and then incubated for 1 hour with polyclonal rabbit anti-CBD antibodies diluted by 1:500 in PBST. The specimens were washed 3X10 minutes in PBST and then incubated for an hour with goat anti rabbit IgG conjugated with 5nm gold particles diluted by 1:100 in PBST. The specimens were washed 2x10 minutes with PBST and once in water. A silver stain kit (BioCell Research Laboratories) was used for final development of the reaction. The specimens were soaked in the combined solutions of the kit for about 10 minutes and then washed with plenty of distilled water. The specimens were observed under light microscope.

13.2. Results

13.2.1. The Effect of CBD on Pollen Tube Growth

Figure 19 shows the effect of CBD compared to BSA (control protein) on pollen tube length. Different letters indicate statistically significant differences between the pollen tube length values ($p \leq 0.05$). It was evident that CBD promoted pollen tube growth. BSA inhibited pollen tube growth probably by osmotic effect. In that respect, the results at 50 μg/ml are the most significant where CBD treated pollen grains produced pollen tubes almost double the

size of that of pollen grains treated with 50 μg/ml BSA.

In an attempt to determine the mode of action of CBD, pollen tubes (CBD treated and control) were stained with white fluorescent brightener (calcofluor) to indicate crystalline cell wall components. Figures 20A and 20B demonstrate that CBD treated pollen tubes produced non-crystalline pollen tube tips as indicated by the absence of the bright color at the tip zone. The absence of rigid crystalline cell wall at the tip zone has the potential to facilitate elongation. Gold-immunolabeling of CBD treated pollen tubes (see Figures 21A and 21B) demonstrated that CBD was bound along the pollen tube but preferably at the tip zone as indicated by the intensive dark dots at the tip zone. Again a strong indication that the absence of rigid crystalline cell wall at the tip zone was the result of CBD. A possible mechanism could be that CBD intercalated with the cellulose fibers while being formed, thereby inhibiting the crystallization process.

13.2.2. The Effect of CBD on Root and Root-Hair Growth

The effect of CBD on root growth of *Arabidopsis thaliana* seedlings was determined, Figure 22. At low concentrations, CBD promoted elongation of roots. At higher concentrations, CBD dramatically reduced root length whereas similar concentration (100 μg/ml) of BSA had no effect. Different letters indicate statistically significant differences between the root length values (≤ 0.05). The effect on shoot length revealed a similar trend however, the differences between the treatments were not dramatic as for the roots, and not statistically significant.

Gold-immunolabeling of CBD treated *Arabidopsis thaliana* seedlings (Figure 23) showed that gold-labeling was restricted to the root and no labeling was shown on the shoot. *Arabidopsis thaliana* seedling staining with white fluorescent brightener (calcofluor) showed that only the root is stained indicating accessible cellulose (Figure 24). The absence of the bright color at the shoot indicates that the stain could not penetrate the waxy cuticle thus explaining why CBD was effective on the roots and not on the shoots.

The inhibitory effect on root elongation by high concentrations (1×10^{-6} to 1×10^{-4} µg/ml) of CBD could be explained by a steric hindrance of the cellulose fibrils thus preventing access of enzymes such as endo-glucanases or xyloglucantransferases or other proteins that modulate cell elongation by loosening the rigid cellulose fibrils network. At low concentrations, CBD could either decrease the crystallinity of the cell wall or alternatively interfere with the xyloglucan cellulose interactions thus preventing crosslinking of the cellulose network.

These results described in this application clearly demonstrate the usefulness of CBD fusion proteins for affinity purifications of proteins and enzymes using cellulose as the insoluble solid matrix. Furthermore, the CBD fusion products offer a wide range of potential applications that would be apparent to one of ordinary skill in view of the above disclosure, including, but not limited to, affinity separation methods and use in diagnostic kits.

14. DEPOSIT OF MICROORGANISMS

The following were deposited with the American Type Culture Collection (ATTC), 12301

Parklawn Drive, Rockville, MD on April 12, 1993, and have been assigned the indicated accession numbers:

Microorganisms	Accession Number
E. coli pCBD-ProA1/2097	69283
E. coli pET-CBD/BL21 (DE3)	69282
E. coli pCBD-ProtA1	75443
E. coli pET-CBD	75444

The invention described and claimed herein is not to be limited in scope by the specific embodiments herein disclosed since these embodiments are intended as illustration of several aspects of the invention. Any equivalent embodiments are intended to be within the scope of this invention. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims.

It is also to be understood that all base pair and amino acid residue numbers and sizes given for nucleotides and peptides are approximate and are used for purposes of description.

A number of references are cited herein, the entire disclosures of which are incorporated herein, in their entirety, by reference.

International Application No: PCT/ /

MICROORGANISMS

Optional Sheet in connection with the microorganism referred to on page 88, lines 1-12 of the description '

A. IDENTIFICATION OF DEPOSIT '

Further deposits are identified on an additional sheet '

Name of depositary institution '

American Type Culture Collection

Address of depositary institution (including postal code and country) '

12301 Parklawn Drive
Rockville, MD 20852
US

Date of deposit ' April 12, 1993 Accession Number ' 69282

B. ADDITIONAL INDICATIONS ' (leave blank if not applicable). This information is continued on a separate attached sheet

C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE ' (if the indications are not all designated States)

D. SEPARATE FURNISHING OF INDICATIONS ' (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later ' (Specify the general nature of the indications e.g., "Accession Number of Deposit")

E. ☑ This sheet was received with the International application when filed (to be checked by the receiving Office)

Virginia L Lilly
(Authorized Officer)

☐ The date of receipt (from the applicant) by the International Bureau "

was

(Authorized Officer)

Form PCT/RO/134 (January 1981)

International Application No: PCT/ /

Form PCT/RO/134 (cont.)

American Type Culture Collection

12301 Parklawn Drive
Rockville, MD 20852
US

Accession No.	Date of Deposit
69283	April 12, 1993
75443	April 12, 1993
75444	April 12, 1993

WHAT IS CLAIMED IS:

1. An isolated cellulose binding domain comprising the amino acid sequence depicted in Figure 1.

2. A cellulose binding domain comprising an amino acid sequence having at least 80% amino acid homology to a contiguous portion of the amino acid sequence depicted in Figure 1.

3. The cellulose binding domain of claim 1 having a binding affinity to cellulose or chitin characterized by a K_d ranging from about 1.5 μM to about 0.5 μM.

4. The cellulose binding domain of claim 3 in which said cellulose or chitin is crystalline.

5. The cellulose binding domain of claim 4 having a K_d of 1.2 μM or less.

6. The cellulose binding domain of claim 4 having a K_d of 1.0 μM or less.

7. The cellulose binding domain of claim 2 having at least about 100 amino acids having greater than 80% amino acid homology to a contiguous portion of the amino acid sequence depicted in Figure 1.

8. The cellulose binding domain of claim 2 having at least about 50 amino acids having greater than 80% amino acid homology to a contiguous portion of the amino acid sequence depicted in Figure 1.

9. The cellulose binding domain of claim 7 having greater than 90% amino acid homology to a contiguous portion of the amino acid sequence depicted in Figure 1.

10. A recombinant nucleic acid encoding a cellulose binding domain comprising the nucleic acid sequence depicted in Figure 1 or its complement.

11. A recombinant nucleic acid encoding a cellulose binding domain comprising a nucleic acid sequence having at least 60% homology to a contiguous portion of the nucleic acid sequence depicted in Figure 1 or its complement.

12. The recombinant nucleic acid of claim 10 which is a DNA polynucleotide.

13. The recombinant nucleic acid of claim 10 which is a RNA polynucleotide.

14. A polymerase chain reaction kit (PCR) comprising a pair of primers capable of priming cDNA synthesis in a PCR reaction, in which each primer is a polynucleotide of claim 12 or its complement.

15. A recombinant nucleic acid comprising a sequence of at least about 300 nucleotides having greater than 60% homology to a contiguous portion of the cellulose binding domain nucleotide sequence depicted in Figure 1 or its complement.

16. A recombinant nucleic acid comprising a sequence of at least about 150 nucleotides having greater than 60% homology to a contiguous portion of

the cellulose binding domain nucleotide sequence depicted in Figure 1 or its complement.

17. A recombinant nucleic acid of claim 15 or 16 having greater than 80% homology to a contiguous portion of the cellulose binding domain nucleotide sequence depicted in Figure 1 or its complement.

18. Plasmid pET-cellulose binding domain as deposited with the ATCC.

19. A recombinant vector comprising a nucleotide sequence encoding a cellulose binding domain of claim 1, 2, 8, 9 or 10.

20. A host cell transfected with a recombinant vector of claim 19.

21. Recombinant *E. coli* pET-cellulose binding domain/BL21(DE3) as deposited with the ATCC having accession No. 69282.

22. A cellulose binding domain fusion protein comprising a cellulose binding domain and a second protein, said cellulose binding domain comprising the amino acid sequence of Figure 1.

23. The cellulose binding domain fusion protein of claim 22 in which the second protein is Protein A, heparinase or an enzyme capable of degrading an environmental pollutant.

24. The cellulose binding domain fusion protein of claim 22 in which the second protein is an

HSP protein, HSP antibody, cross-reactive HSP-related protein or peptide, or an antigenic protion thereof.

25. The cellulose binding domain fusion protein of claim 22 in which said second protein is a recombinant antibody.

26. The cellulose binding domain fusion protein of 25 comprising cellulose binding domain-V_L-V_H.

27. The cellulose binding domain fusion protein of claim 22 in which said second protein is a hormone.

28. The cellulose binding domain fusion protein of claim 22 in which said second protein is an enzyme.

29. The cellulose binding domain fusion protein of claim 22 in which said second protein is selected from the group consisting of Protein A, Protein G, streptavidin, avidin, Taq polymerase, non-Taq polymerases, alkaline phosphatase, RNase, DNase, restriction enzymes, peroxidases, glucanases, chitinases, beta and alfa glucosidases, beta and alpha glucoronidases, amylase, transferases, beta-lactamase, non-beta lactamase antibiotic modifying and degrading enzymes, luciferase, esterases, lipases, proteases, bacteriocines, antibiotics, enzyme inhibitors, growth factors, hormones, receptors, membranal proteins, nuclear proteins, transcriptional and translational factors and nucleic acid modifying enzymes.

30. Plasmid pcellulose binding domain-ProtA1 as deposited with the ATCC.

31. A recombinant vector comprising a nucleotide sequence encoding a cellulose binding domain fusion protein of claim 22 or 23.

32. Recombinant *E. coli* pcellulose binding domain-ProtA1/2097 as deposited with the ATCC, having Accession No. 75443.

33. A method for the production of a cellulose binding domain fusion product comprising:

(a) providing nucleic acid encoding the cellulose binding domain fusion product, said cellulose binding domain fusion product comprising a cellulose binding domain and a second protein, said cellulose binding domain being capable of binding cellulose or chitin with a high affinity and being substantially free of other proteins with which it is naturally associated, and said second protein comprising N-terminal and C-terminal ends;

(b) transfecting a host cell with said nucleic acid; and

(c) culturing the transformed host cell under conditions suitable for the expression of said cellulose binding domain fusion product.

34. The method of claim 33 in which said nucleic acid further encodes for a cleavage site upstream of the N-terminal amino acid of said second protein of said cellulose binding domain fusion product.

35. The method of claim 33 in which said nucleic acid further encodes for a cleavage site downstream of the C-terminal amino acid of said second protein of said cellulose binding domain fusion product.

36. The method of claim 33 in which said nucleic acid further encodes for a first cleavage site upstream of said N-terminal amino acid of said second protein of said cellulose binding domain fusion product and a second cleavage site downstream of said C-terminal amino acid of said second protein of said cellulose binding domain fusion product.

37. A method of purifying a cellulose binding domain fusion product comprising:

(a) contacting a mixture comprising a recombinant cellulose binding domain fusion product with an effective amount of cellulose under conditions suitable for the formation of an insoluble binding complex comprising cellulose and said recombinant cellulose binding domain fusion product;

(b) isolating said insoluble cellulose-cellulose binding domain fusion product binding complex; and

(c) recovering said cellulose binding domain fusion product from said insoluble cellulose-cellulose binding domain fusion product binding complex.

38. The method of claim 37 in which said cellulose is replaced by chitin.

39. The method of claim 37 or 38 in which said recovery comprises exposing said insoluble

cellulose-cellulose binding domain fusion product binding complex to a releasing reagent and isolating the released cellulose binding domain fusion product from said cellulose or chitin.

40. The method of claim 39 in which said releasing reagent comprises 6 M urea or 6 M guanidine-HCl.

41. The method of claim 34, 35 or 36 which further comprises exposing said cellulose binding domain fusion product to a cleavage agent capable of cleaving said first or second cleavage site or both.

42. The method of claim 41 in which said cleavage agent is an enzyme.

43. The method of claim 41 in which said cleavage agent is a chemical cleavage agent.

44. The host cell of claim 20 which is a bacterial host cell.

45. The host cell of claim 20 which is a mammalian host cell.

46. A diagnostic kit for the detection of a substance of interest comprising:

(a) a cellulose binding domain fusion product comprising (i) a cellulose binding domain capable of binding to cellulose with high affinity and substantially free of other proteins with which it is naturally associated, and (ii) a second protein capable binding a substance of interest;

(b) a detectable label; and

(c) cellulose.

47. The diagnostic kit of claim 46 in which said cellulose is replaced by chitin.

48. The diagnostic kit of claim 46 or 47 in which said second protein is Protein A, or an enzyme.

49. The diagnostic kit of claim 46 or 47 in which said second protein is an HSP protein, HSP antibody, HSP-related protein, peptide or antigenic portion thereof.

50. The diagnostic kit of claim 46 or 47 in which said cellulose binding domain fusion product further comprises a ligand affinity bound to said second protein, said ligand capable of binding said substance of interest.

51. The diagnostic kit of claim 50 in which said ligand is an antibody.

52. An immunoassay method of detecting the presence of a substance of interest in a test sample comprising:

(a) incubating a test sample, which may contain a substance of interest, with a sufficient amount of a cellulose binding domain fusion product comprising (i) a cellulose binding domain capable of binding to cellulose with high affinity and substantially free of other proteins with which it is naturally associated, and (ii) a second protein capable of binding said substance of interest, under conditions that allow for the binding of said

substance of interest to said second protein of said cellulose binding domain fusion product;

(b) adding an amount of cellulose effective to bind the amount of said cellulose binding domain fusion product used in step (a) to provide an insoluble cellulose-cellulose binding domain fusion product binding complex;

(c) separating said insoluble cellulose-cellulose binding domain fusion product binding complex from unbound components;

(d) incubating said insoluble cellulose-cellulose binding domain fusion product binding complex with a sufficient amount of a detectable label, said label capable of binding to said substance of interest; and

(e) separating said insoluble cellulose-cellulose binding domain fusion product binding complex of step (d) from unbound components and determining the presence or absence of said label, to provide an indication of the presence or absence of said substance of interest in said test sample.

53. A method of detecting the presence of a substance of interest in a test sample comprising:

(a) contacting a test sample, which may contain a substance of interest, with an insoluble matrix capable of immobilizing said substance of interest;

(b) incubating said insoluble matrix with a sufficient amount of a cellulose binding domain fusion product comprising (i) a cellulose binding domain capable of binding to cellulose with high affinity and substantially free of other proteins with which it is naturally associated, and (ii) a second protein capable of binding said immobilized substance

of interest, under conditions that allow for the binding of said immobilized substance of interest to said second protein of said cellulose binding domain fusion product;

(c) separating said insoluble matrix of step (b) from unbound components;

(d) incubating said insoluble matrix of step (c) with a detectable label capable of binding said substance of interest or said cellulose binding domain fusion product under conditions that allow for the binding of said label to said substance of interest or said cellulose binding domain fusion product; and

(e) separating said insoluble matrix of step (d) from unbound components and determining the presence or absence of said label, to provide an indication of the presence or absence of said substance of interest in said test sample.

54. The method of claim 53 which further comprises (i) contacting said insoluble matrix of step (c) with a sufficient amount of cellulose under conditions that allow for the binding of said cellulose to said cellulose binding domain fusion product to form a cellulose-cellulose binding domain fusion product binding complex, and (ii) separating said insoluble matrix of step (i) from unbound components, including unbound cellulose.

55. The method of claim 54 in which said label is capable of binding said substance of interest or said cellulose-cellulose binding domain fusion protein binding complex.

56. The method of claim 55 in which said label is capable of binding to the cellulose of said cellulose-cellulose binding domain fusion protein binding complex.

57. The method of claim 52, 53 or 55 in which said test sample comprises a bodily fluid.

58. The method of claim 52 or 54 in which said cellulose is replaced by chitin.

59. The method of claim 53 or 54 in which said insoluble matrix is an electrophoresis gel blot.

60. The method of claim 52, 53 or 54 in which said substance of interest is a protein or peptide and said second protein is an antibody against said protein or peptide.

61. The method of claim 52, 53 or 54 in which said substance of interest is an antibody and said second protein is Protein A.

62. The method of claim 52 or 54 in which said substance of interest comprises a biotinylated probe bound to a protein, peptide, hormone, nucleic acid or other probe-targetable molecule and said second protein is streptavidin.

63. The method of claim 52, 53 or 54 in which said label comprises a radioisotope, a fluorescent molecule or an enzyme.

64. The method of claim 56 in which said label comprises a cellulose binding domain fusion

product comprising cellulose binding domain and alkaline phosphatase.

65. The method of claim 56 in which said label comprises a cellulose binding domain fusion product comprising cellulose binding domain and horse radish peroxidase.

66. The method of claim 64 or 65 which further comprises adding a sufficient amount of a substrate for said enzyme, which substrate is converted by said enzyme to a detectable compound.

67. The method of claim 52 in which step (d) is performed by incubating said insoluble cellulose-cellulose binding domain fusion product binding complex with a sufficient amount of a detectable label comprising a labeled substance of interest, said label capable of binding to any second protein of said cellulose binding domain fusion product which remains unbound to said substance of interest, and in which step (e) is performed by separating said insoluble cellulose-cellulose binding domain fusion product binding complex of step (d) from unbound components and comparing the signal observed from said test sample relative to the signal observed from a control sample.

68. The method of claim 52 in which said cellulose binding domain fusion product is included in a dip stick.

69. A dip stick useful in detecting a substance of interest in a test sample comprising a cellulose binding domain fusion product, said

cellulose binding domain fusion product comprising (i) a cellulose binding domain capable of binding to cellulose with high affinity and substantially free of other proteins with which it is naturally associated, and (ii) a second protein capable of binding a substance of interest.

70. The dip stick of claim 69 in which said second protein is selected from the group consisting of Protein A, HSP protein, HSP antibody, cross-reactive HSP-related protein or peptide or an antigenic portion thereof, an enzyme, hormone, antigen, and antibody.

71. A signal amplification system comprising:

(a) a first cellulose binding domain fusion product comprising (i) a cellulose binding domain capable of binding to cellulose with high affinity and substantially free of other proteins with which it is naturally associated, and (ii) a second protein capable of binding a chimeric probe, said probe further capable of binding a substance of interest; and

(b) a second cellulose binding domain fusion product comprising (i) a cellulose binding domain capable of binding to cellulose with high affinity and substantially free of other proteins with which it is naturally associated, and (ii) an enzyme capable of acting on a substrate to produce a detectable compound.

72. A signal amplification system comprising:

(a) a cellulose binding domain fusion product comprising (i) a cellulose binding domain capable of binding to cellulose with high affinity and substantially free of other proteins with which it is naturally associated, and (ii) a second protein capable of binding a chimeric probe, said probe further capable of binding a substance of interest; and

(b) labeled cellulose binding domain, said cellulose binding domain retaining its capacity to bind to cellulose with high affinity and substantially free of other proteins with which it is naturally associated.

73. The signal amplification system of claim 71 or 72 which further comprises said probe.

74. The signal amplification system of claim 71 or 72 which further comprises pebble-milled cellulose.

75. The signal amplification system of claim 71 or 72 in which said probe is biotinylated and said second protein is streptavidin.

76. The signal amplification system of claim 71 or 72 in which said enzyme is selected from the group consisting of alkaline phosphatase and horse radish peroxidase.

77. The signal amplification system of claim 72 in which said label comprises a radioisotope, a fluorescent molecule or an enzyme.

78. A drug delivery system comprising cellulose binding domain associated with a drug, said cellulose binding domain retaining its capacity to bind to cellulose with high affinity and substantially free of other proteins with which it is naturally associated.

79. The drug delivery system of claim 80 in which said drug is conjugated to said cellulose binding domain either directly or through a linker moiety.

80. The drug delivery system of claim 78 in which said drug is an antifungal agent.

81. The drug delivery system of claim 78 in which said drug is selected from the group consisting of Amphotericin B, Nystatin, Undecylenic Acid, and Clotrimazole.

82. The drug delivery system of claim 78 in which said drug is an imidazole.

83. The drug delivery system of claim 78 which may be administered parenterally, orally, topically or by inhalation.

84. The drug delivery system of claim 78 which may be administered intranasally, opthalmically or intravaginally.

85. The drug delivery system of claim 78 which is in the form of a solid, gel, liquid or aerosol.

86. The drug delivery system of claim 78 in which said drug is effective against a disease or infection caused by a yeast or fungal agent whose cellular membrane contains a cellulosic or chitinic substance.

87. The drug delivery system of claim 86 in which said agent is *Aspergillus fumigatus*, a member of the genus *Candida* or *Monilia*, or an epidermatocyte.

```
          10          20          30          40          50
          .           .           .           .           .
GCA GCG ACA TCA TCA ATG TCA GTT GAA TTT TAC AAC TCT AAC AAA TCA GCA CAA
CGT CGC TGT AGT AGT TAC AGT CAA CTT AAA ATG TTG AGA TTG TTT AGT CGT GTT
Ala Ala Thr Ser Ser Met Ser Val Glu Phe Tyr Asn Ser Asn Lys Ser Ala Gln>
                                    10

     60          70          80          90          100
     .           .           .           .           .
ACA AAC TCA ATT ACA CCA ATA ATC AAA ATT ACT AAC ACA TCT GAC AGT GAT TTA
TGT TTG AGT TAA TGT GGT TAT TAG TTT TAA TGA TTG TGT AGA CTG TCA CTA AAT
Thr Asn Ser Ile Thr Pro Ile Ile Lys Ile Thr Asn Thr Ser Asp Ser Asp Leu>
     20                                     30

110         120         130         140         150         160
.           .           .           .           .           .
AAT TTA AAT GAC GTA AAA GTT AGA TAT TAT TAC ACA AGT GAT GGT ACA CAA GGA
TTA AAT TTA CTG CAT TTT CAA TCT ATA ATA ATG TGT TCA CTA CCA TGT GTT CCT
Asn Leu Asn Asp Val Lys Val Arg Tyr Tyr Tyr Thr Ser Asp Gly Thr Gln Gly>
          40                                     50

     170         180         190         200         210
     .           .           .           .           .
CAA ACT TTC TGG TGT GAC CAT GCT GGT GCA TTA TTA GGA AAT AGC TAT GTT GAT
GTT TGA AAG ACC ACA CTG GTA CGA CCA CGT AAT AAT CCT TTA TCG ATA CAA CTA
Gln Thr Phe Trp Cys Asp His Ala Gly Ala Leu Leu Gly Asn Ser Tyr Val Asp>
                    60                                     70

220         230         240         250         260         270
.           .           .           .           .           .
AAC ACT AGC AAA GTG ACA GCA AAC TTC GTT AAA GAA ACA GCA AGC CCA ACA TCA
TTG TGA TCG TTT CAC TGT CGT TTG AAG CAA TTT CTT TGT CGT TCG GGT TGT AGT
Asn Thr Ser Lys Val Thr Ala Asn Phe Val Lys Glu Thr Ala Ser Pro Thr Ser>
                          80                                     90
```

FIG.1A

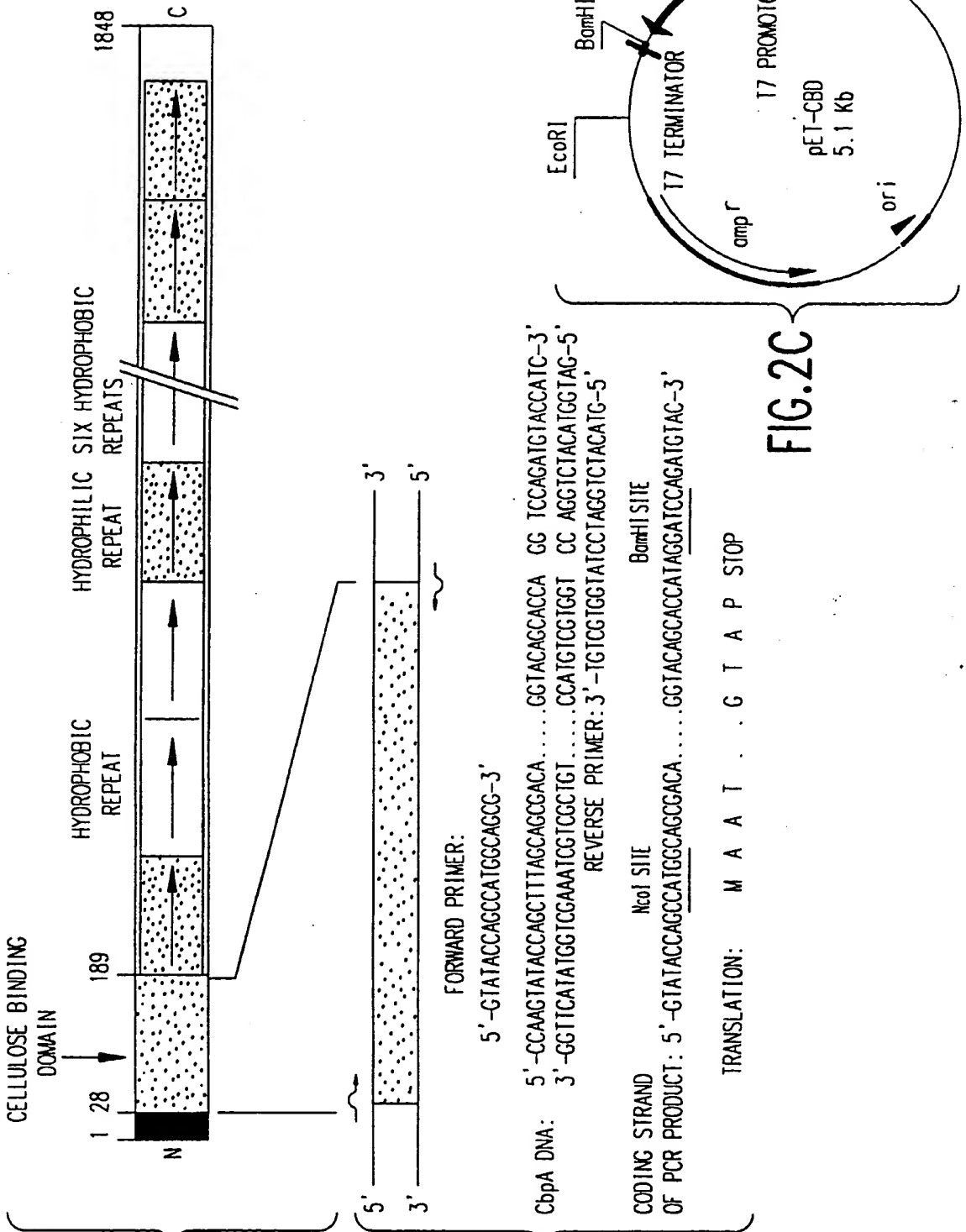
```
          280         290         300         310         320
          .           .           .           .           .
ACC TAT GAT ACA TAT GTT GAA TTT GGA TTT GCA AGC GGA CGA GCT ACT CTT AAA
TGG ATA CTA TGT ATA CAA CTT AAA CCT AAA CGT TCG CCT GCT CGA TGA GAA TTT
Thr Tyr Asp Thr Tyr Val Glu Phe Gly Phe Ala Ser Gly Arg Ala Thr Leu Lys>
                                    100

    330         340         350         360         370
    .           .           .           .           .
AAA GGA CAA TTT ATA ACT ATT CAA GGA AGA ATA ACA AAA TCA GAC TGG TCA AAC
TTT CCT GTT AAA TAT TGA TAA GTT CCT TCT TAT TGT TTT AGT CTG ACC AGT TTG
Lys Gly Gln Phe Ile Thr Ile Gln Gly Arg Ile Thr Lys Ser Asp Trp Ser Asn>
    110                                 120

380         390         400         410         420         430
.           .           .           .           .           .
TAC ACT CAA ACA AAT GAC TAT TCA TTT GAT GCA AGT AGT TCA ACA CCA GTT GTA
ATG TGA GTT TGT TTA CTG ATA AGT AAA CTA CGT TCA TCA AGT TGT GGT CAA CAT
Tyr Thr Gln Thr Asn Asp Tyr Ser Phe Asp Ala Ser Ser Ser Thr Pro Val Val>
            130                                 140

        440         450         460         470         480
        .           .           .           .           .
AAT CCA AAA GTT ACA GGA TAT ATA GGT GGA GCT AAA GTA CTT GGT ACA GCA CCA
TTA GGT TTT CAA TGT CCT ATA TAT CCA CCT CGA TTT CAT GAA CCA TGT CGT GGT
Asn Pro Lys Val Thr Gly Tyr Ile Gly Gly Ala Lys Val Leu Gly Thr Ala Pro>
                    150                                 160
```

FIG.1B



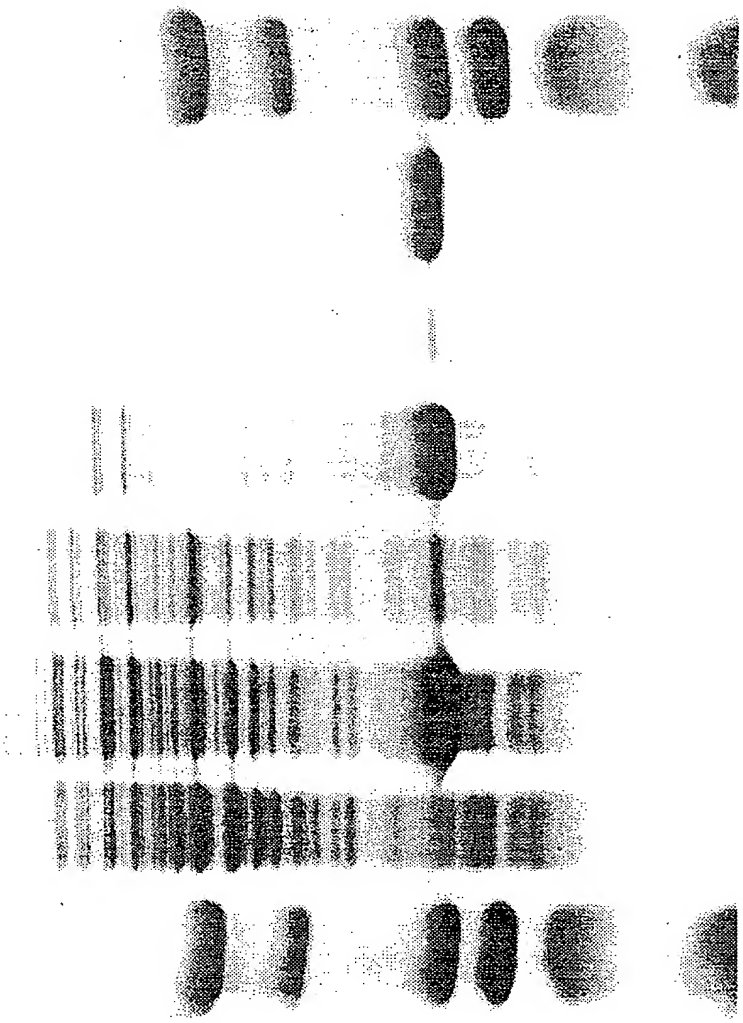


FIG. 3

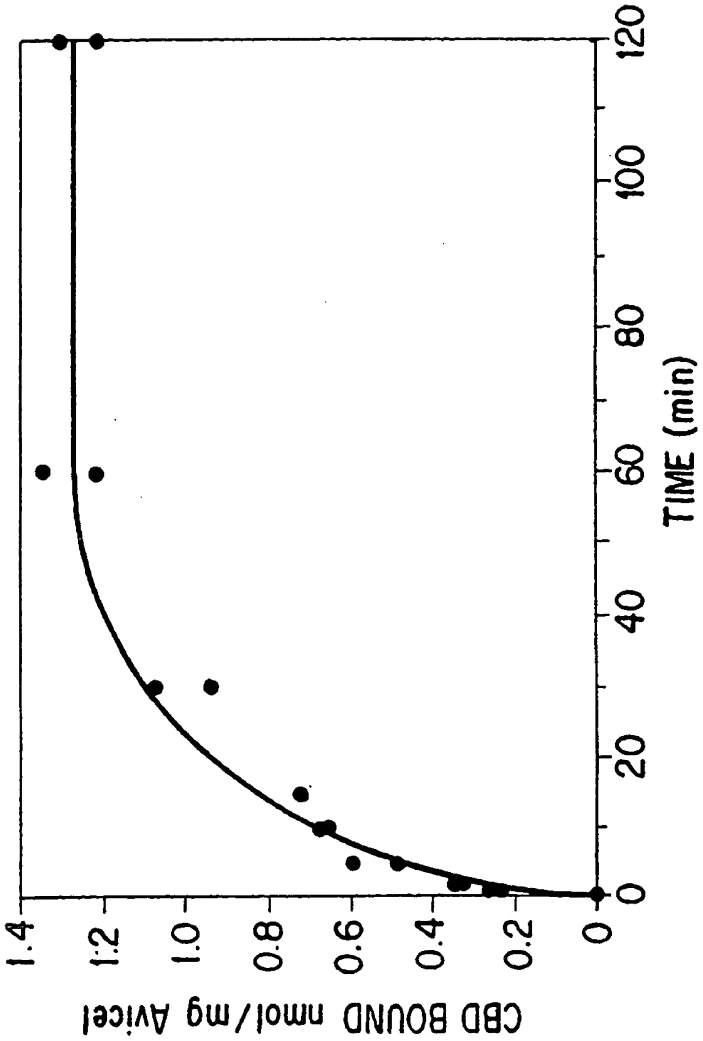


FIG. 4

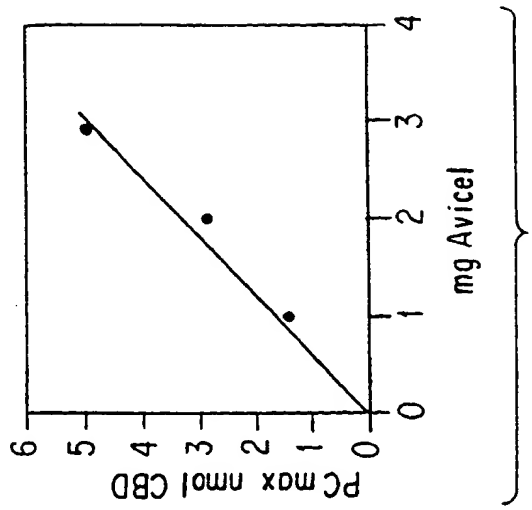


FIG. 5B

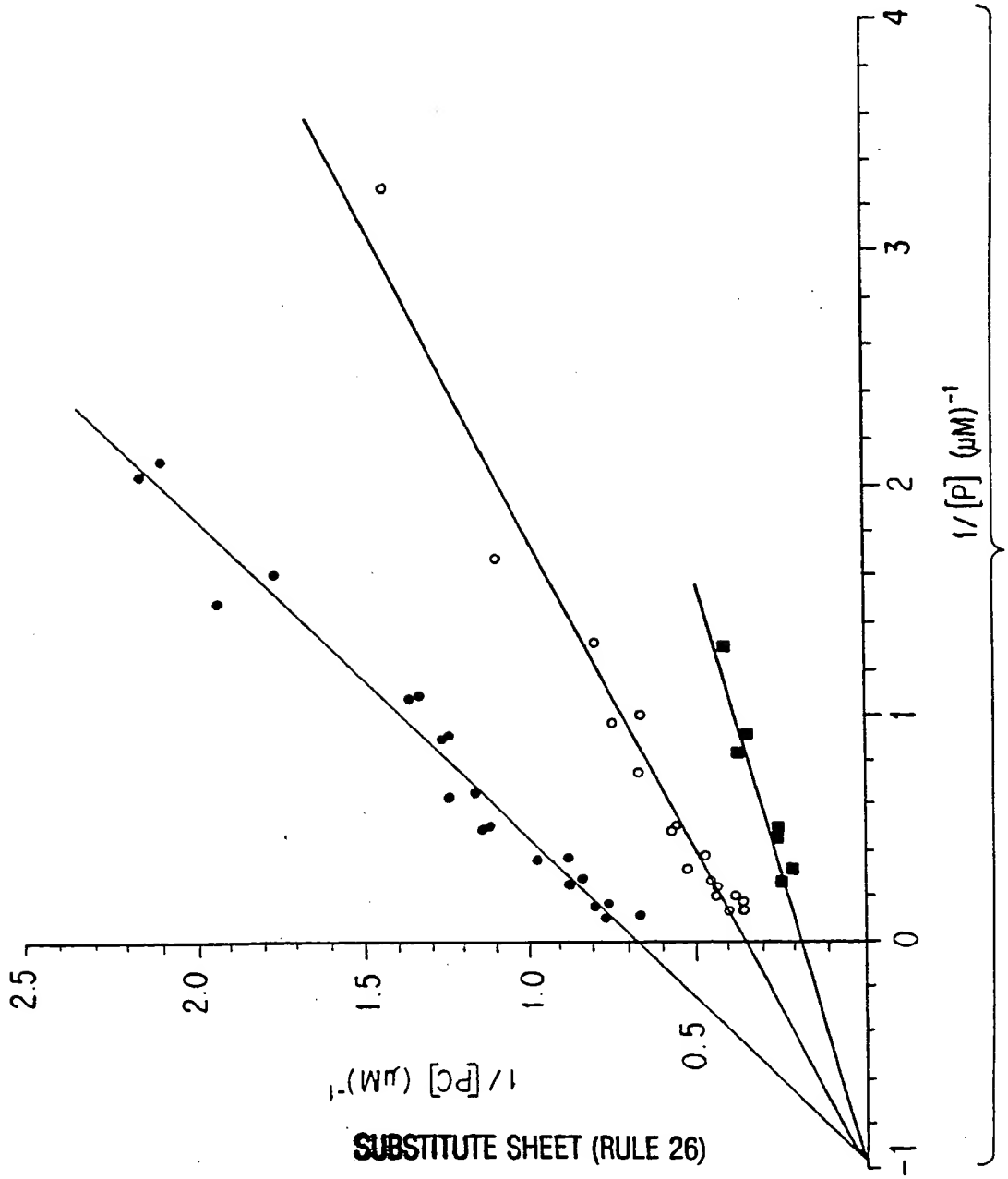


FIG. 5A

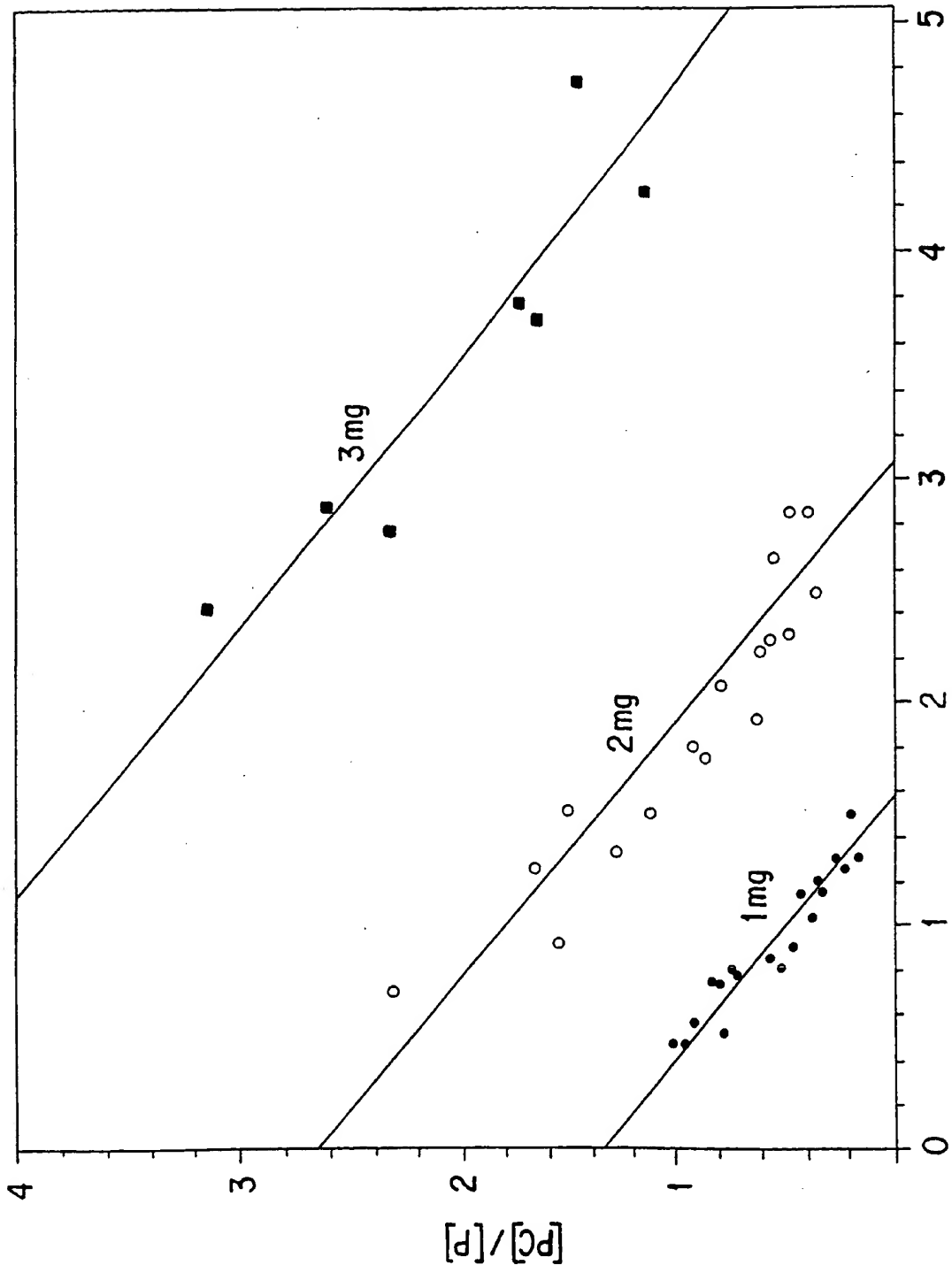


FIG.6

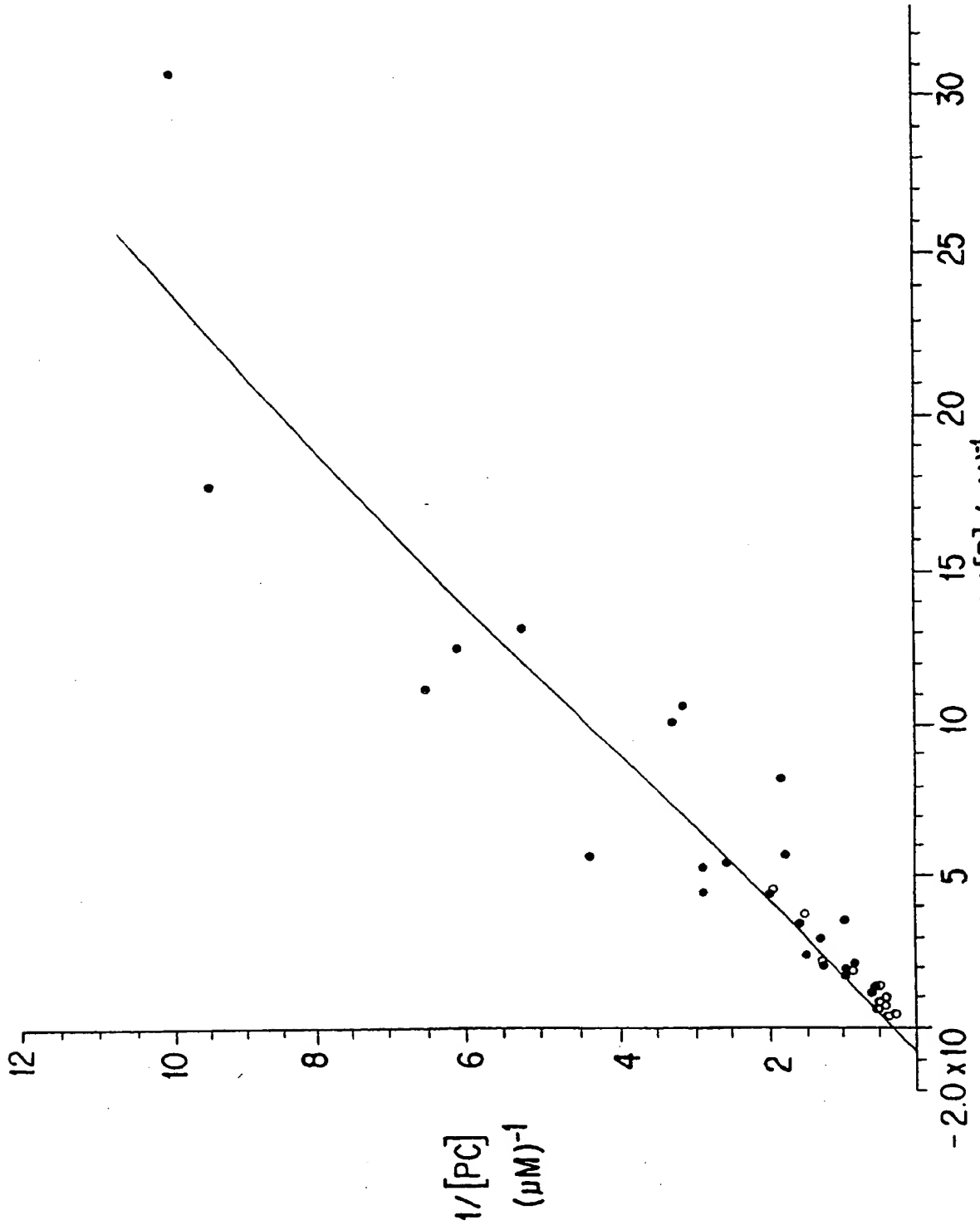


FIG.7

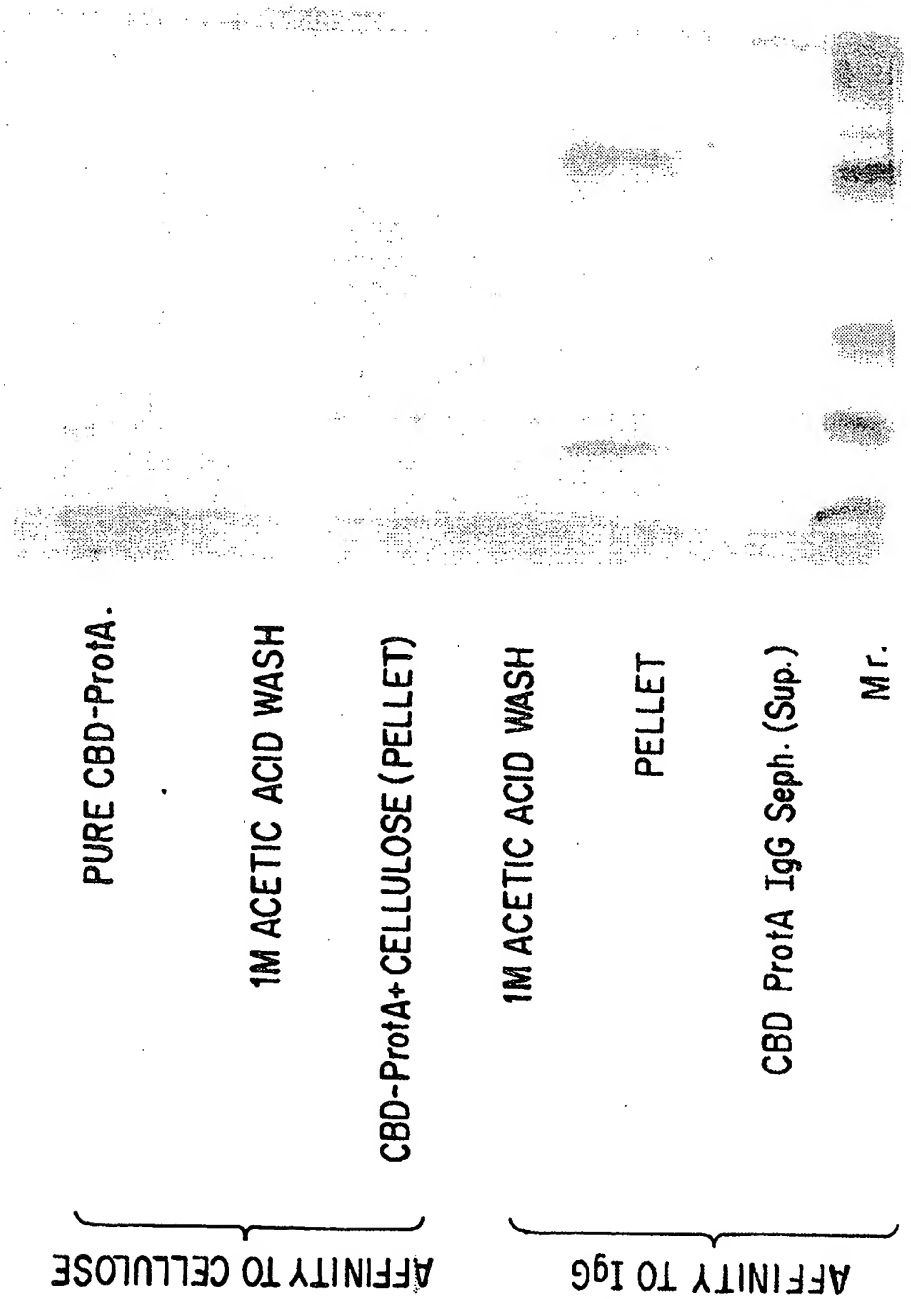


FIG.8

SUBSTRATE:	OBSERVED K_d (in μM)	OBSERVED PC_{max} (in μmoles CBD/g)
AVICEL PH101	1.1 ± 0.3	1.4 ± 0.3
SIGMACELL 20	1.1	1.2
SIGMACELL 50	1.4	1.7
SIGMACELL 100	1.3	0.5
MICROGRANULAR CELLULOSE	1.0	0.4
FIBROUS CELLULOSE	1.4	0.2
COTTON	0.8	6.4
CELLULON	1.2	5.3
XYLAN	–	0
SEPHADEX G-75	–	0
NIGERAN	–	0
CHITIN	1.0	1.6
AVICEL PH101		
+CMC (4mg/ml)	1.1	1.1
+CELLOBIOSE (4mg/ml)	1.0	1.5

FIG.9

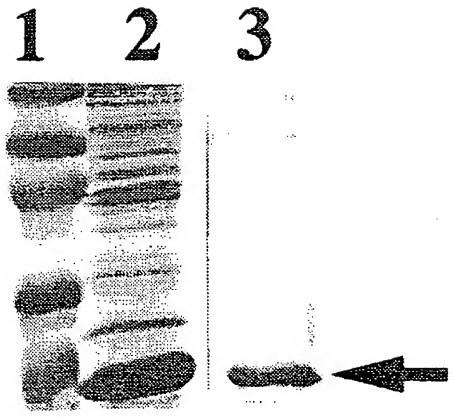


FIG. 10

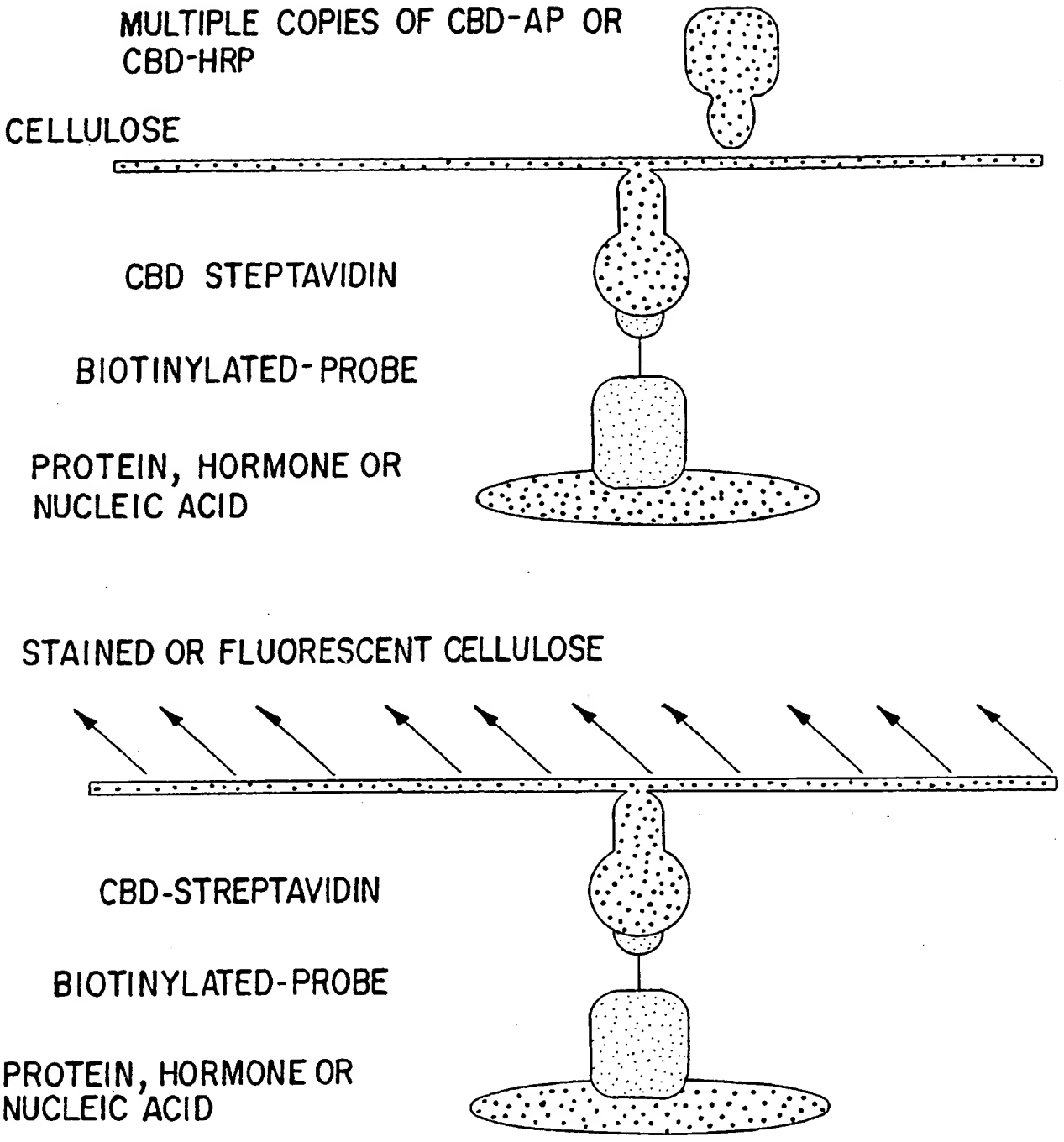


FIG. 11

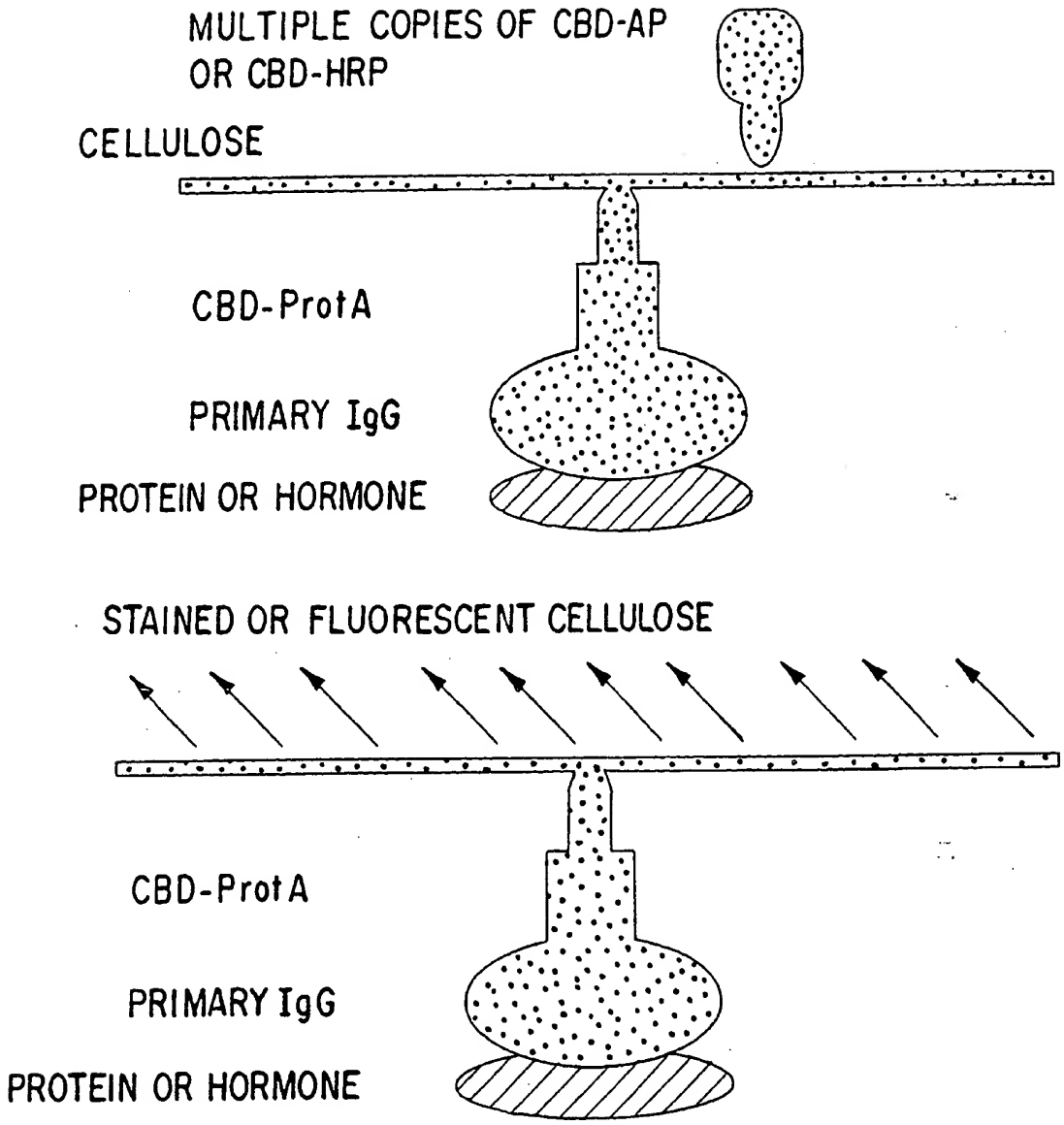


FIG. 12

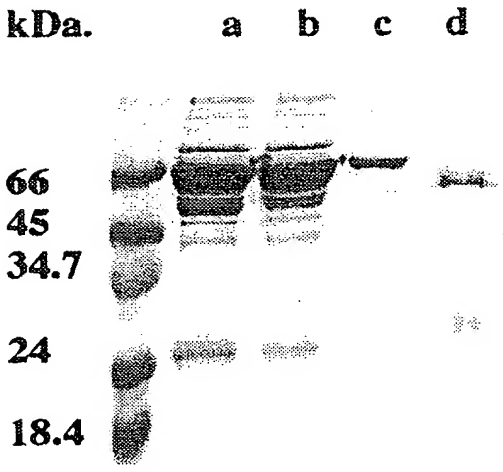


FIG. 13

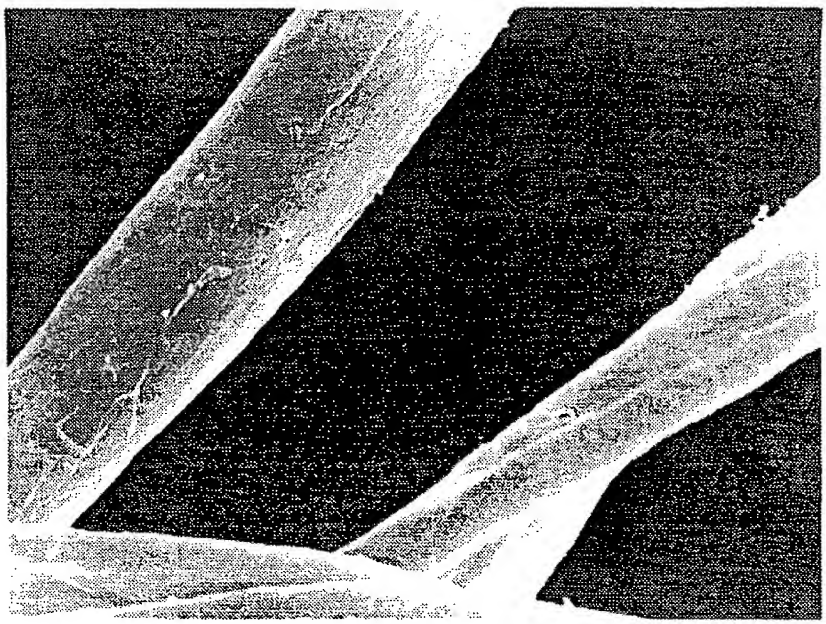


FIG. 14A

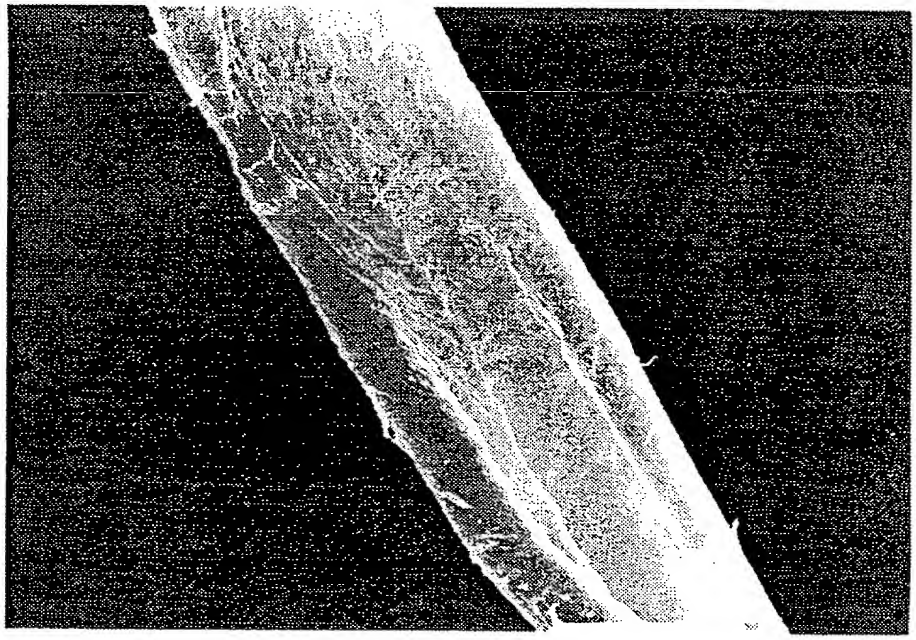


FIG. 14B

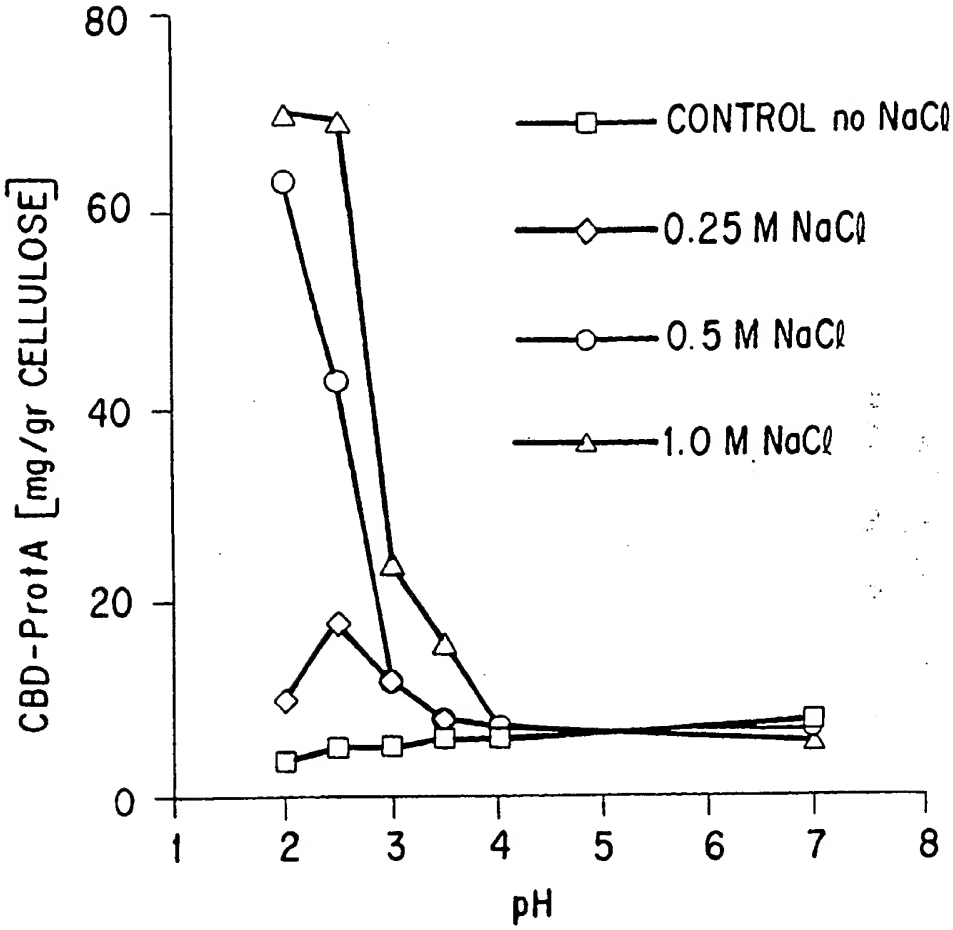


FIG.15

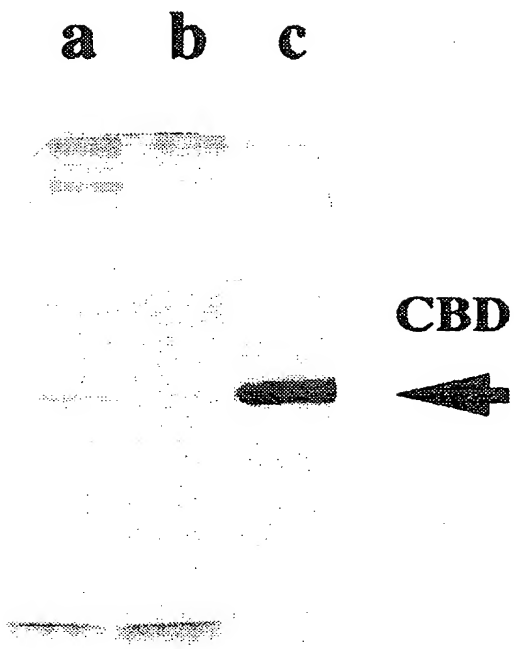


FIG. 16

a b

CBD

FIG. 17

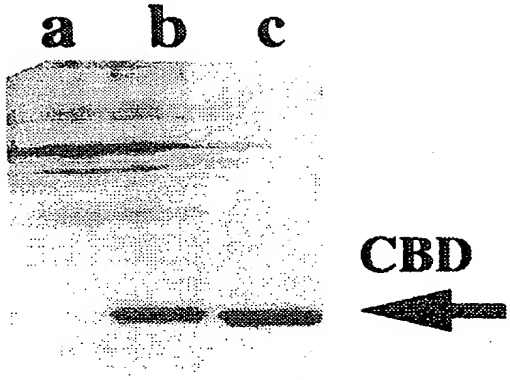


FIG. 18

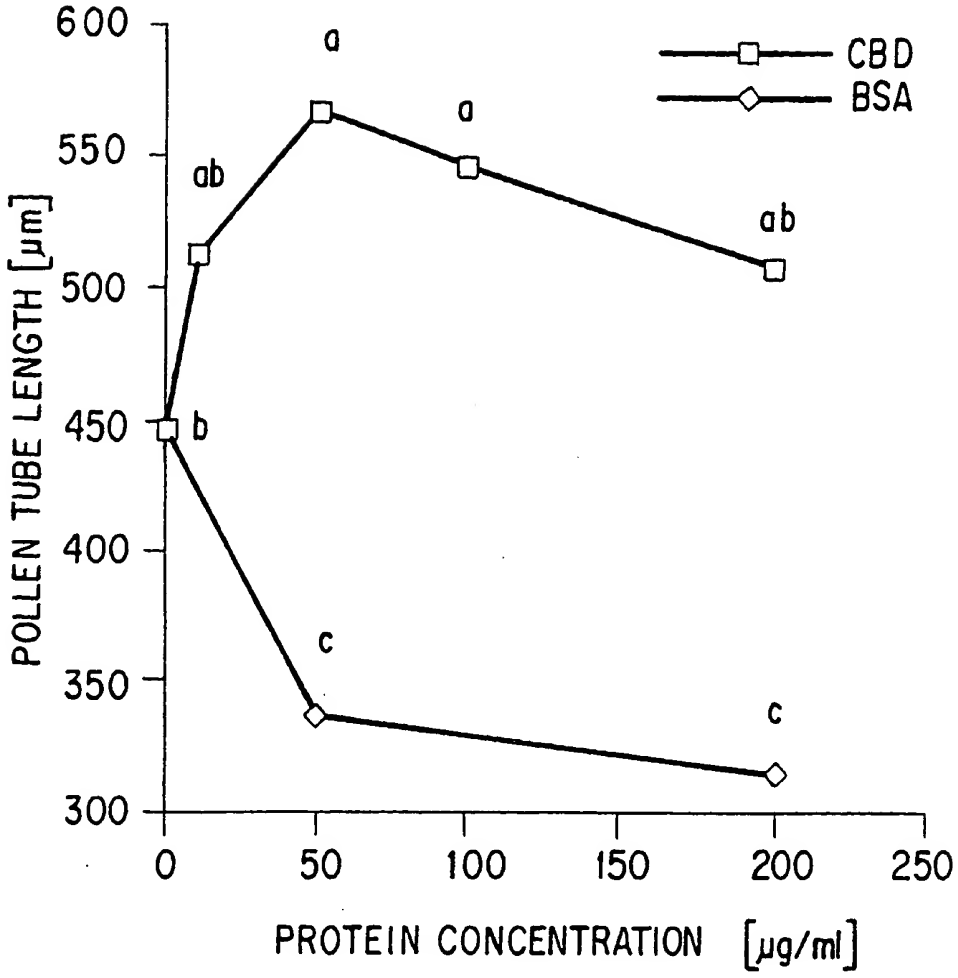

600
550
500
450
400
350
300
POLLEN TUBE LENGTH [μm]
0
50
100
150
200
250
PROTEIN CONCENTRATION [μg/ml]
CBD
BSA
a
a
ab
ab
b
c
c

FIG. 19

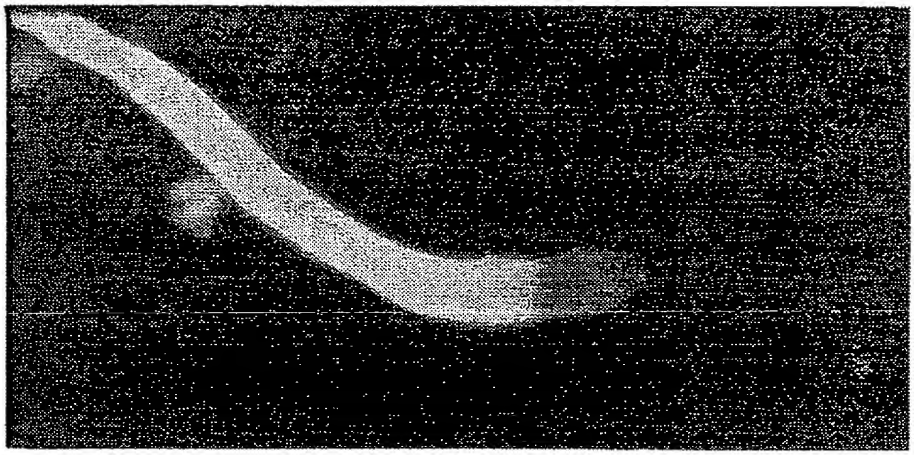


FIG. 20A

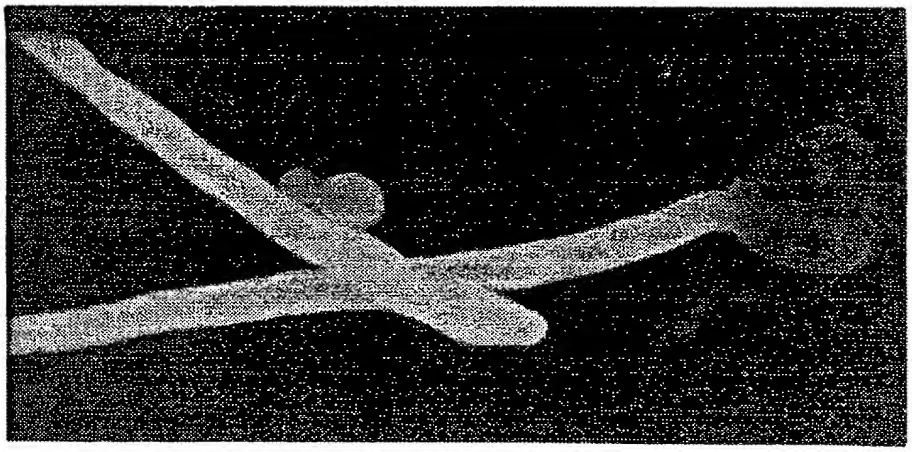


FIG. 20B

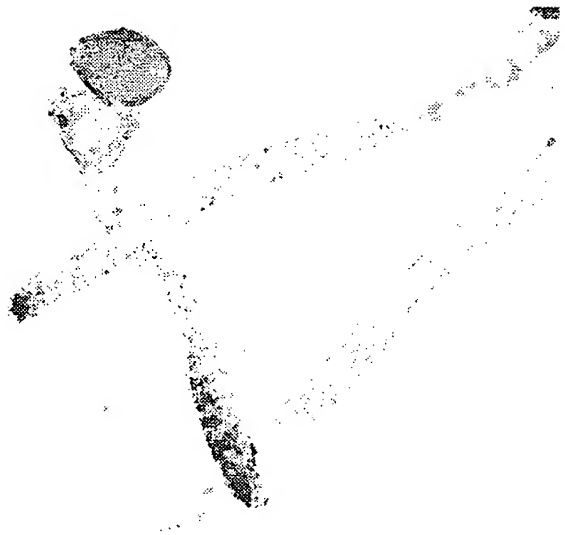


FIG. 21A

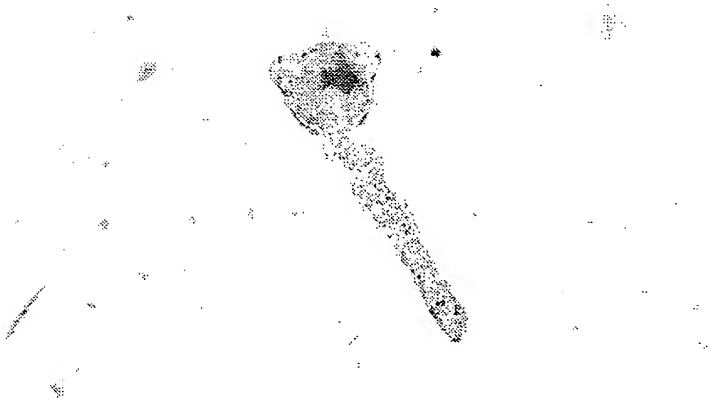


FIG. 21B

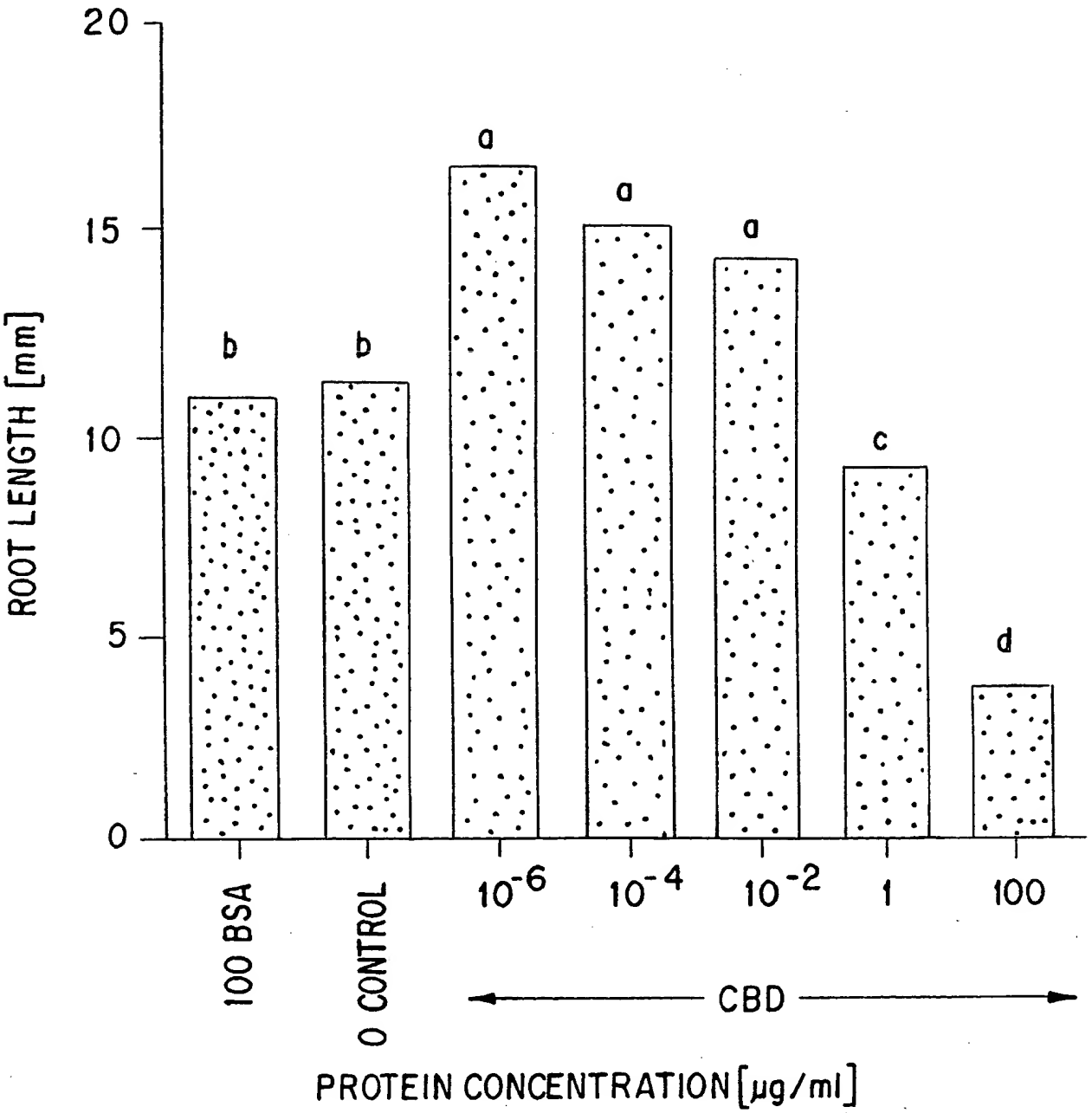


FIG. 22

FIG. 23

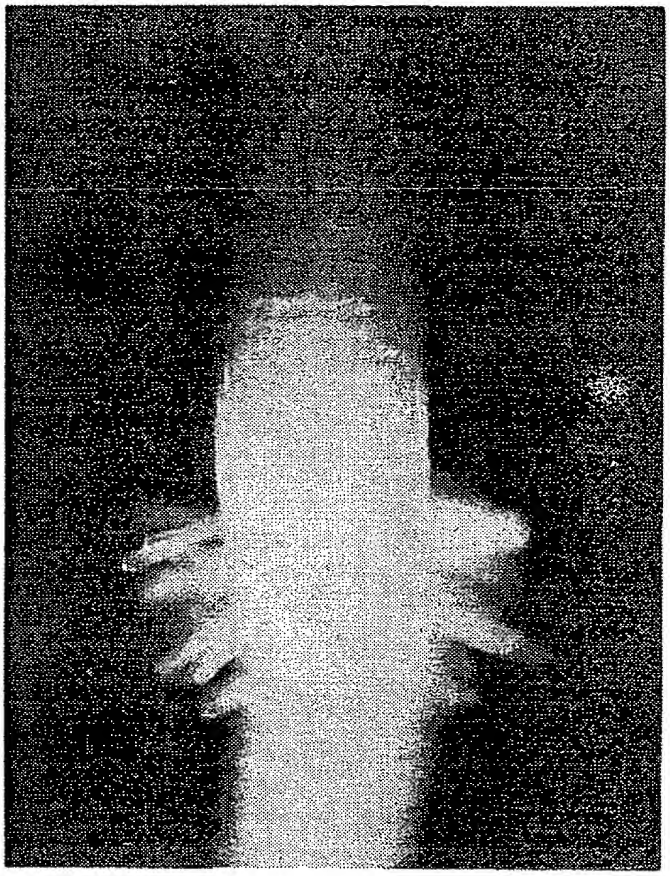


FIG.24

INTERNATIONAL SEARCH REPORT

Int ational application No.
PCT/US94/04132

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/350; 536/23.1, 23.4, 23.7, 24.3, 24.33; 435/69.1, 172.1, 252.3, 320.1; 436/501, 512; 514/2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS; DIALOG; SEQUENCE DATABASES INCLUDING GENBANK AND EMBL

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ---- Y	US, A, 5,137,819 (KILBURN ET AL) 11 August 1992, see entire document.	2-9, 11-13, 15-17, 19, 20, 21, 33-37, 39-41, 44, ---- 14, 22-29, 31, 38, 42, 43, 45--87
X ---- Y	FEMS Microbiology Letters, Volume 99, issued 1992, Poole et al, "Identification of the cellulose-binding domain of the cellulosome subunit S1 from Clostridium thermocellum YS", pages 181-186, see entire document.	2, 11-13, 15, 16, 19, 20 ---- 3-9, 14, 17

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search	Date of mailing of the international search report
14 JULY 1994	22 JUL 1994

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	ELIZABETH C. KEMMERER
Facsimile N . (703) 305-3230	Telephone N . (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/04132

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ---- Y	Molecular and General Genetics, Volume 223, issued 1990, Jauris et al, "Sequence analysis of the Clostridium stercorarium celZ gene encoding a thermoactive cellulase (Avicelase I): Identification of catalytic and cellulose-binding domains", pages 258-267, see entire document.	2, 11-13, 15, 16, 19, 20 ---- 3-9, 11-17, 19, 20
Y, P	US, A, 5,229,501 (KEIFER ET AL) 20 July 1993, see entire document.	14
Y	US, A, 5,122,448 (VAUGHAN ET AL) 16 June 1992, see entire document.	46-77
Y,E	US, A, 5,322,769 (BOLLING ET AL) 21 June 1994, see entire document.	46-77
Y	Biotechnology and Bioengineering, Volume 42, issued 1993, Ong et al, "The cellulose-binding domain (CBDCex) of an exoglucanase from Cellulomonas fimi: production in Escherichia coli and characterization of the polypeptide", pages 401-409, see entire document.	3-9, 11-17, 19, 20
Y	Biochemistry Journal, Volume 273, issued 1991, Durrant et al, "The non-catalytic C-terminal region of endoglucanase E from Clostridium thermocellum contains a cellulose-binding domain", pages 289-293, see entire document.	3-9, 11-17, 19, 20
Y	TIBTECH, Volume 7, issued September 1989, Ong et al, "The cellulose -binding domains of cellulases: tools for biotechnology", pages 239-243, see entire document.	1-89
Y	Biochemistry Journal, Volume 279, issued 1991, Poole et al, "Characterization of hybrid proteins consisting of the catalytic domains of Clostridium and Ruminococcus endoglucanases, fused to Pseudomonas non-catalytic cellulose-binding domains", pages 787-792, see entire document.	22-32
Y	FEBS Letters, Volume 244, Number 1, issued February 1989, Greenwood et al, "Fusion to an endoglucanase allows alkaline phosphatase to bind to cellulose", pages 127-131, see entire document.	22-46

INTERNATIONAL SEARCH REPORT

In ational application No.
PCT/US94/04132

Box I Observations where certain claims were f und unsearchable (Continuation f item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ N protest accompanied the payment of additional search fees.

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (5):

C07K 13/00; C07H 17/00; C12P 21/06; C12N 15/00, 1/20; G01N 33/566, 33/563; A61K 37/00

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

530/350; 536/23.1, 23.4, 23.7, 24.3, 24.33; 435/69.1, 172.1, 252.3, 320.1; 436/501, 512; 514/2

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

I. Claims 1-21, 44, and 45, drawn to an isolated cellulose binding domain peptide, DNA encoding same, a PCR kit, plasmids and host cells.
II. Claims 22-36, drawn to a cellulose binding domain fusion protein, plasmids, vectors, host cells, and a method of producing the fusion protein.
III. Claims 37-43, drawn to a method of purifying a cellulose binding domain fusion protein.
IV. Claims 46-70, drawn to a diagnostic kit and immunoassay.
V. Claims 71-77, drawn to a signal amplification system.
VI. Claims 78-87, drawn to a drug delivery system.
Inventions I-VI represent separate inventive concepts. Invention I is directed to the general inventive concept of the recombinant production of a cellulose binding domain peptide, which is useful for the production of antibodies, for example. Invention II is directed to the general inventive concept of the recombinant production of a fusion protein comprising a cellulose binding domain, which is useful for degrading environmental pollutants (as claimed in claim 23), for example. Invention III is directed to the general inventive concept of methods of purifying fusion products comprising the cellulose binding domain, which requires the consideration of purification techniques such as chromatography, which is not required by any of the other inventions. Invention IV is directed to the general inventive concept of diagnostic kits and immunoassays, requiring consideration of binding kinetics and detection sensitivity which is not required by any of the other inventions. Invention V isdirected to the general inventive concept of a signal amplification system, which is useful for in situ hybridization assays. Finally, invention VI is directed to the general inventive concept of a drug delivery system, requiring consideration of disease states, efficacy, toxicity, and conventional drugs, which is not required by any of the other groups. The claims are not so linked by a special technical feature within the meaning of PCT Rule 13.2 so as to form a single inventive concept.